



SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Zsebo, Krisztina M.
Bosselman, Robert A.
Suggs, Sidney V.
Martin, Francis H.
- (ii) TITLE OF INVENTION: Stem Cell Factor
- (iii) NUMBER OF SEQUENCES: 104
- (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
(B) STREET: 6300 Sears Tower, 233 South Wacker Drive
(C) CITY: Chicago
(D) STATE: Illinois
(E) COUNTRY: United States of America
(F) ZIP: 60606-6402
- (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: 08/449,649
(B) FILING DATE: 24-MAY-1995
(C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 07/982,255
(B) FILING DATE: 25-NOV-1992
- (vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 07/589,701
(B) FILING DATE: 01-OCT-1990
- (vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 07/573,616
(B) FILING DATE: 24-AUG-1990
- (vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 07/537,198
(B) FILING DATE: 11-JUN-1990
- (vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 07/422,383
(B) FILING DATE: 16-OCT-1989
- (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Clough, David W.
(B) REGISTRATION NUMBER: 36,107
(C) REFERENCE/DOCKET NUMBER: 01017/32953
- (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 312/474-6300
(B) TELEFAX: 312/474-0448
(C) TELEX: 25-3856

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 165 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Glu Glu Ile Cys Arg Asn Pro Val Thr Asp Asn Val Lys Asp Ile Thr
1 5 10 15
Lys Leu Val Ala Asn Leu Pro Asn Asp Tyr Met Ile Thr Leu Asn Tyr
20 25 30
Val Ala Gly Met Asp Val Leu Pro Ser His Cys Trp Leu Arg Asp Met
35 40 45
Val Thr His Leu Ser Val Ser Leu Thr Thr Leu Leu Asp Lys Phe Ser
50 55 60
Asn Ile Ser Glu Gly Leu Ser Asn Tyr Ser Ile Ile Asp Lys Leu Gly
65 70 75 80
Lys Ile Val Asp Asp Leu Val Ala Cys Met Glu Glu Asn Ala Pro Lys
85 90 95
Asn Val Lys Glu Ser Leu Lys Lys Pro Glu Thr Arg Asn Phe Thr Pro
100 105 110
Glu Glu Phe Phe Ser Ile Phe Asn Arg Ser Ile Asp Ala Phe Lys Asp
115 120 125
Phe Met Val Ala Ser Asp Thr Ser Asp Cys Val Leu Ser Ser Thr Leu
130 135 140
Gly Pro Glu Lys Asp Ser Arg Val Ser Val Thr Lys Pro Phe Met Leu
145 150 155 160
Pro Pro Val Ala Ala
165

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: modified_base
- (B) LOCATION: 9
- (D) OTHER INFORMATION: /mod_base= Inosine

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ACRTTYTTNG GNGCRTTYTC YTCCAT

26

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: modified_base
- (B) LOCATION: 12 & 15
- (D) OTHER INFORMATION: /mod_base= Inosine

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AARAAATCYT CNGGNGTRAA RTT

23

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 14 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GTATCNGGYT TYTT

14

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATGGARGARA AYGCCCCCAA RAAYGT

26

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CCNAAYGAYT AYATGWTMAC

20

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 20 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GGNGGNARCA TRAANGGYTT 20

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 23 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

ACCAKAARAT CTTYAAANCG ATC 23

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 22 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GTATTTTCAA TAGATCCATT GA 22

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 14 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CCAACTATGT CGCC 14

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 21 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:
 GTAGTCAAGC TGACTGATAA G 21

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 21 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:
 TAACCAACAA TGACTAGGCA A 21

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 16 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:
 TTCCAGAGTC AGTGTC 16

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 29 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:
 GCGAAGCTTG CCTTCCTTA TGAAGAAGA 29

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 38 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:
 GCGCCGCGGT TACGGTGGTA ACATGAAGGG CTTGTGA 38

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GATAAATGCA AGTGATAATC C

21

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 36 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GCGGTCGACC CGCGGAACTT TAAGTCCATG CAACAC

36

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 36 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CACCCGCGGT TATGCAACAG GGGGTAACAT AAATGG

36

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 36 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CACCCGCGGT TAGGCTGCAA CAGGGGGTAA CATAAA

36

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 18 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:
CTTAATGTTG AAGAAACC 18

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 22 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:
GATGGTAGTA CAATTGTCAG AC 22

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 22 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:
GTCTGACAAT TGTACTACCA TC 22

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 22 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:
CAATTTAGTG ACGTCTTTTA CA 22

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 24 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:
TTAGATGAGT TTTCTTTCAC GCAC 24

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

AAATCATTCA AGAGCCCAGA ACCC

24

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

AACATCCATC CCGGGGAC

18

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

CTGGCAATAT TTTAAGTCTC AAGAAGACC

29

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GCGCCGCGGC TCCTATAGGT GCTAATTGG

29

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:
CCTCACC ACT GTTTGTGCTG GATCGCA 27

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:
GGTGTCTAGA CTTGTGTCTT CTCATAAGG A 31

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:
CCCCCCHGG 10

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:
TTTTTTTTTT TTTTTTTTGG 20

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:
TTTTTTTTTT TTTTTTTTAG 20

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

TTTTTTTTTT TTTTTTTTCG

20

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

TTCGGCCGAT CAGGCCCCC CCCC

24

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

TTCGGCCGGA TAGGCCTTTT TTTTTTTTTT

30

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

GGCCGGATAG GCCTCACNNN NNNT

24

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

GGCCGGATAG GCCTCAC

17

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4673 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: join(660..773, 1184..1246, 2053..2223, 2837..2993, 3692..3774)

(ix) FEATURE:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION: join(720..773, 1184..1246, 2053..2223, 2837..2993, 3692..3774)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

| | |
|---|-----|
| AAAGTATCTT TCTATTGGCG AAGGACATGT TTTCCCATAA GTGGTAAACA AACTGTCTGC | 60 |
| ACATAATAAT TATCTTGCTG CCGTAAAGAT TAGGTTAAAT TCTGCCTTCG ATCTAAAAAC | 120 |
| ACACCCTTCT GTCAATCCGA GGAGCAGTGT GCTAGTCTAG AGGTCTAAAT GAAGGCTCCT | 180 |
| TTCACGGTTG TATTTCTGCT CCCCAAATTG TCCACATTTA AAAGGAGAGT GCTTCTTTTC | 240 |
| AGCCTTAGGC TCTGAATTTT ATGCATTCCT CCATTTTCCG AGGTCCCCC CAAGTGATAA | 300 |
| TTCTGTTACA CGTTGCTACA AGTTCATCCC TAATTGCCGT CAAGAACTG ACTGTAGAAG | 360 |
| GCTTACCACA GACGTTGTAA CCGACAGTAA AGCCATTGAA AGAGTAATTC AAACAGGATG | 420 |
| GAAGCCAGGA GTATTTTGTG GCTGTTGCTC TTTTCTTTT CAGTTTGGTG AGAGCAGCTT | 480 |
| GAATGCTTAA CATTTAAGCC ATCAGCTTAA AACAAAACAA AACAAAACAA AAAAAACCC | 540 |
| CGCTCTGGCA TATTTGCACT TAACACATAC GGTATAAGGT GTTACTGGTT TGCATAGTTC | 600 |
| TGGATTTTTT TTTTTTAAAA ACTGATGGAC ACCAAGAAAT GTTTCTGTTC TTTGTTTAG | 659 |
| ACT TGG ATT ATC ACT TGC ATT TAT CTT CAA CTG CTC CTA TTT AAT CCT | 707 |
| Thr Trp Ile Ile Thr Cys Ile Tyr Leu Gln Leu Leu Leu Phe Asn Pro | |
| -20 -15 -10 -5 | |
| CTC GTC AAA ACT CAG GAG ATC TGC AGG AAT CCT GTG ACT GAT AAT GTA | 755 |
| Leu Val Lys Thr Gln Glu Ile Cys Arg Asn Pro Val Thr Asp Asn Val | |
| 1 5 10 | |
| AAA GAC ATT ACA AAA CTG GTAAGTAAAG AATGATTTTG GCATCTATAA | 803 |
| Lys Asp Ile Thr Lys Leu | |
| 15 | |
| GTCTTCCCTG TGCTTGCTGA CCACATAGGT TCAGGGCACT CCCGACAGGA GTTCCCAGCT | 863 |
| TTCTAAGATA AGGAATCACT GTACGAGTCT GAAGTGCTTC TTCTGGGCAA ATGGGAGATG | 923 |

| | |
|--|------|
| CTTAGGTCAT GGAGGGTTTA TCTGTATAAC TGGCCCTTTG CACACCAACA AAGTGA CTGA | 983 |
| CTGGCTTTTG CCTGTTACCT ACTGTCTCCA GTCCTGGGCA TGGTATATAC TTAGGCACCC | 1043 |
| AAGATTGGAT TTACAACTCA AGCATTATAT ATTGGACAAC ACGGGGTATG AGATATTAAT | 1103 |
| GATATGTCAG GTTGGATGGA TGAGTTTTCT CAAGAAATTC TCTTGATTTT ACTCACGTTT | 1163 |
| TCATTTCTTG GTCTCTGTAG GTG GCG AAT CTT CCA AAT GAC TAT ATG ATA | 1213 |
| Val Ala Asn Leu Pro Asn Asp Tyr Met Ile | |
| 20 25 | |
| ACC CTC AAC TAT GTC GCC GGG ATG GAT GTT TTG GTATGTAGTC CACACACTTC | 1266 |
| Thr Leu Asn Tyr Val Ala Gly Met Asp Val Leu | |
| 30 35 | |
| TGAGTTGCCT TTTAGTAGCT AATGGGTGAC CTGTGCTTAT TCACATTGAA GACATTATTT | 1326 |
| GCTCTTTGTC GTTTTTAGAT GTTGACCTAT AATTTTTCCT TCAAGCTGCT GCTAAGATTA | 1386 |
| TCAGTGAGCA TTTCAGTATG TGTTTTAAGC CTA CTCATTA AAAGGAAATG GCTCATCTTA | 1446 |
| GACGTAGCAA CCGATGTTAA TTTTTCCTCA GGCATCTCTC AGAGGGACTT GAATGTTAA | 1506 |
| ATCATGTTAA ATTTCTCCT TGGCTATGTT ATTTCTCATG GCTATGTTAT TCCTATTCGT | 1566 |
| ATTTCAITTA AAGGGACGGA ATATTTATTG TATTTCTGAA CTTTTTCAGG CATGCATCCG | 1626 |
| GGTCTTTGAA TAAAACACTA AGACTCCTTC TAGTAATGTT TGTAACTCTG TCTGTATCGA | 1686 |
| ATGTCTTTGA AAACGCAGTG ACTAAGCCAT AAATAATCTT CCACAGAACG TCCAGTG GTT | 1746 |
| CATGAAC TTT GTATGTGGGG GTGGGGCAAG AATTGTCTCA CTATTGGTCA AGGAAGAGAA | 1806 |
| GGTAAGGTAT GCAAGGGTGG TTTAATCTTC TTCCGTGGAA GGACAAAATC ATCTATCATT | 1866 |
| TCCTCTGATC TCTATGCATT TGTTTGTTTT GAACTGAATC TGACTTGAGC AAGAGTTGGC | 1926 |
| GTCCTGTGTT CTGAGGAAAC TCTTTGTCTT GCAGTCAGTG ACTAAAAGTG CTGAGAGATC | 1986 |
| TGAAGAGCAC TCTGAATCTG CCATATTTTT AATAGATGCT TTGTCTTCTC TTTGAATTTT | 2046 |
| TTCCAG CCT AGT CAT TGT TGG TTA CGA GAT ATG GTA ACA CAC TTA TCA | 2094 |
| Pro Ser His Cys Trp Leu Arg Asp Met Val Thr His Leu Ser | |
| 40 45 50 | |
| GTC AGC TTG ACT ACT CTT CTG GAC AAG TTT TCA AAT ATT TCT GAA GGC | 2142 |
| Val Ser Leu Thr Thr Leu Leu Asp Lys Phe Ser Asn Ile Ser Glu Gly | |
| 55 60 65 | |
| TTG AGT AAT TAT TCC ATC ATA GAC AAA CTT GGG AAA ATA GTG GAT GAC | 2190 |
| Leu Ser Asn Tyr Ser Ile Ile Asp Lys Leu Gly Lys Ile Val Asp Asp | |
| 70 75 80 85 | |
| CTC GTG GCA TGT ATG GAA GAA AAT GCA CCT AAG GTA ACTTGGT ATTCATCAGA | 2243 |
| Leu Val Ala Cys Met Glu Glu Asn Ala Pro Lys | |
| 90 95 | |
| ATTATTTTTT TTATACTGAG CTCATGATGA GCAATTCACA ACCACTTGTA ATTCCAGCTC | 2303 |
| CAGAGGACAT TATCCCCTCT TTGGATGCCA TAGGAATCTG CTCTCAAATA TG TAGATACC | 2363 |
| ACCTCTGCCA CCTCAGCACA TACATACACA TAATTAAAAA ATAGAAACAT TAAAGGAGTT | 2423 |

| | |
|--|------|
| CCAATCAATC CTTATTCTTT TCTGTATTCA GSTATGCCAG ATGTAAATTC TAGGAATATG | 2483 |
| TTTTAAAGGC TAATTCTTAT TTTGTAATAA GCAGCTTTAA AATTCTTAAT TGTTTTTTTCG | 2543 |
| GGTCACTTTA TTGTCCTATT GCCACGACAT TGTCTGTGCC CATTGTCTGT TATTCCTTCT | 2603 |
| GTTTTGTTTA TTGTTCCCTA GTTACTTTGA TCATGAGATT GACCTGTTAC CCGTTGTTAT | 2663 |
| TCTCTGTAGC CATTTTGAGT TGTGTCTATT AGAACAGCTG TTAAATTACT TGAATCATTG | 2723 |
| AGGACATAGT CAATAATCTA TTATGCTGAT CCAGTCAAGT CTATGAGTTA TTTGAAAAC | 2783 |
| AGAATCTTTG TTAATTATTT GTTTGCTTGT TTGTTTGTTT ATTATTTGTC TAG AAT | 2839 |
| | Asn |
| GTA AAA GAA TCA CTG AAG AAG CCA GAA ACT AGA AAC TTT ACT CCT GAA | 2887 |
| Val Lys Glu Ser Leu Lys Lys Pro Glu Thr Arg Asn Phe Thr Pro Glu | |
| 100 105 110 | |
| GAA TTC TTT AGT ATT TTC AAT AGA TCC ATT GAT GCC TTC AAG GAC TTC | 2935 |
| Glu Phe Phe Ser Ile Phe Asn Arg Ser Ile Asp Ala Phe Lys Asp Phe | |
| 115 120 125 | |
| ATG GTG GCA TCT GAC ACT AGT GAT TGT GTG CTC TCT TCA ACA TTA GGT | 2983 |
| Met Val Ala Ser Asp Thr Ser Asp Cys Val Leu Ser Ser Thr Leu Gly | |
| 130 135 140 145 | |
| CCT GAG AAA G GTAAGGCTTT TAAGCATTTT TTGTTTAAAT GTACATAGAA | 3033 |
| Pro Glu Lys | |
| AGCCTGAACT TCTGTAAGCC TCTACTGCTG AATCAACTAA ATGTGTTGCT GTAGAAAGAA | 3093 |
| CGTGTGGGTT TTTCTGATAA AAACAAAAAG CAAATATCAA TGACTACCAA TGATTATTAT | 3153 |
| CTAGCTTGAG AGATATGCCC TAAGACAGCG ATTCTCGATA TTTCTAAATT AAAGAATTGT | 3213 |
| GTGATGGTGG CTCACATATT TTCTAACTGT GATATTTGCC AGGAGAGTAG AATAATGTTA | 3273 |
| TTCTTCATCC CCAGAATTCC TAAGATTTCA CGTCTCATGT CTTTTCCATA AGGTTCAAAC | 3333 |
| TCTGAGACTT GAGTTCTGAG CCTCAGCAGG TCATTCTGAA TCCCCACTCT CCCCAGAGCTG | 3393 |
| GGTCCCTATG GGGGAACTAA CTTCAATTGCT TTCTTTTAAA ACATGACGAG TTACCAACAG | 3453 |
| CTCCTCGCTA TTATAAACAT GTTCCTAAGC ATGTCTGTGC ATGCAATAAG CCTTCACTCT | 3513 |
| ACAAGACAGT TATGGTGTAT CGCTTGACAA AACTGAGCAG CCAAGCTGAG TATGAAATAA | 3573 |
| TAATCTAGAC TTGGGAGGCA GACCCAGCAC CTACTGTGAT ATTGCACTTC GCCTTTGGGG | 3633 |
| GACTCTATGA TTCAAAAGTT CACCATGTAA CACTGACACA TTATTGCTTT CTATTTAG AT | 3693 |
| | Asp |
| TCC AGA GTC AGT GTC AGA AAA CCA TTT ATG TTA CCC CCT GTT GCA GCC | 3741 |
| Ser Arg Val Ser Val Arg Lys Pro Phe Met Leu Pro Pro Val Ala Ala | |
| 150 155 160 165 | |
| AGT TCC CTT AGG AAT GAC AGC AGT AGC AGT AAT AGTAAGTACA CATATCTGAT | 3794 |
| Ser Ser Leu Arg Asn Asp Ser Ser Ser Ser Asn | |
| 170 175 | |

| | |
|---|------|
| TTACTGCATG CATGGCTCCA AGTATCCTCT ATAGGAGTGT TGCATGGACT TAAAGTTTAT | 3854 |
| AAATCACTAC TAATAATGCT GTTCTGTCAC TGTATTTCCT TGTATGGGCT TCCTGACAAT | 3914 |
| TAAATATCTG GTTTGTAGAA TCGGATCTCC TTAGAGGTTA AGATGACCAT GACAAAATTA | 3974 |
| GGCCAATCAA CTTTCTGCGA AGGTTATTTT AAATAAGGCA CGAAATTAAT TGAAGGAAAA | 4034 |
| AAAAATACAA GCAAGGCCTT ATTTTGAATC ATGGTAGGCT TAAAATAGAC TTTGTGGAGA | 4094 |
| ATGTCCCTGA TCAAAGTGGA GTTTTCAGAT TTCAAGTGCA TGTGCTAACT CTCCACAATG | 4154 |
| TCAAGGCTAT TTTCAGTTTT GTGTCTCCAT ATTTACTACT GCATGTTTGG AAATTTGCTG | 4214 |
| ATGCTGTTAG ATTACCTAAG AATGTATGTT GAAGAAGAAT GGACTTCTTT CCCTAAAATT | 4274 |
| TCTGTCCTCT TTGCCCAAGA ACCCACGTTT CTGGAAGACT ATCTTATTTT CATGTCTGTG | 4334 |
| CAATGATCAT TATAAAGATT ATTGAATATA CTGGAATAC TCTGGTTTCT GTTTTTACAG | 4394 |
| ATTCATAATA GCTTATTCAG TCTTTAAAGA AAGTTCTCTG AAGTCCATGC TTTAGAATGT | 4454 |
| TTCTCTATCA AAAGTTGACC TGGACCTTAA ATAAAGCTAT ATTTAGTCTT TTTATCCCTG | 4514 |
| AAAAATATAT TTCACAGTGT AGACATTTGA TATACATCTA AGGGAAGGAT GCTGCCAGAA | 4574 |
| TGCTCGGGCT GGCAGTCTAC AAAGTCCACT GCTCTCAGGA TGGACTTCTG AAAGCGGAAA | 4634 |
| TTGTGAACTG CATGCATATA ACATATCAGA TCCTCGAGC | 4673 |

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 196 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

| | |
|---|-------------|
| Thr Trp Ile Ile Thr Cys Ile Tyr Leu Gln Leu Leu Leu Phe Asn Pro | |
| -20 | -15 -10 -5 |
| Leu Val Lys Thr Gln Glu Ile Cys Arg Asn Pro Val Thr Asp Asn Val | |
| | 1 5 10 |
| Lys Asp Ile Thr Lys Leu Val Ala Asn Leu Pro Asn Asp Tyr Met Ile | |
| | 15 20 25 |
| Thr Leu Asn Tyr Val Ala Gly Met Asp Val Leu Pro Ser His Cys Trp | |
| | 30 35 40 |
| Leu Arg Asp Met Val Thr His Leu Ser Val Ser Leu Thr Thr Leu Leu | |
| | 45 50 55 60 |
| Asp Lys Phe Ser Asn Ile Ser Glu Gly Leu Ser Asn Tyr Ser Ile Ile | |
| | 65 70 75 |
| Asp Lys Leu Gly Lys Ile Val Asp Asp Leu Val Ala Cys Met Glu Glu | |
| | 80 85 90 |

Asn Ala Pro Lys Asn Val Lys Glu Ser Leu Lys Lys Pro Glu Thr Arg
 95 100 105
 Asn Phe Thr Pro Glu Glu Phe Phe Ser Ile Phe Asn Arg Ser Ile Asp
 110 115 120
 Ala Phe Lys Asp Phe Met Val Ala Ser Asp Thr Ser Asp Cys Val Leu
 125 130 135 140
 Ser Ser Thr Leu Gly Pro Glu Lys Asp Ser Arg Val Ser Val Arg Lys
 145 150 155
 Pro Phe Met Leu Pro Pro Val Ala Ala Ser Ser Leu Arg Asn Asp Ser
 160 165 170
 Ser Ser Ser Asn
 175

(2) INFORMATION FOR SEQ ID NO:41:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 849 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 26..844

(ix) FEATURE:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION: 101..844

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

| | |
|---|-----|
| CTGGATCGCA GCGCTGCCTT TCCTT ATG AAG AAG ACA CAA ACT TGG ATT ATC | 52 |
| Met Lys Lys Thr Gln Thr Trp Ile Ile | |
| -25 -20 | |
| ACT TGC ATT TAT CTT CAA CTG CTC CTA TTT AAT CCT CTC GTC AAA ACT | 100 |
| Thr Cys Ile Tyr Leu Gln Leu Leu Leu Phe Asn Pro Leu Val Lys Thr | |
| -15 -10 -5 | |
| CAG GAG ATC TGC AGG AAT CCT GTG ACT GAT AAT GTA AAA GAC ATT ACA | 148 |
| Gln Glu Ile Cys Arg Asn Pro Val Thr Asp Asn Val Lys Asp Ile Thr | |
| 1 5 10 15 | |
| AAA CTG GTG GCG AAT CTT CCA AAT GAC TAT ATG ATA ACC CTC AAC TAT | 196 |
| Lys Leu Val Ala Asn Leu Pro Asn Asp Tyr Met Ile Thr Leu Asn Tyr | |
| 20 25 30 | |
| GTC GCC GGG ATG GAT GTT TTG CCT AGT CAT TGT TGG TTA CGA GAT ATG | 244 |
| Val Ala Gly Met Asp Val Leu Pro Ser His Cys Trp Leu Arg Asp Met | |
| 35 40 45 | |
| GTA ACA CAC TTA TCA GTC AGC TTG ACT ACT CTT CTG GAC AAG TTT TCA | 292 |
| Val Thr His Leu Ser Val Ser Leu Thr Thr Leu Leu Asp Lys Phe Ser | |
| 50 55 60 | |

| | |
|---|-----|
| AAT ATT TCT GAA GGC TTG AGT AAT TAT TCC ATC ATA GAC AAA CTT GGG | 340 |
| Asn Ile Ser Glu Gly Leu Ser Asn Tyr Ser Ile Ile Asp Lys Leu Gly | |
| 65 70 75 80 | |
| AAA ATA GTG GAT GAC CTC GTG GCA TGT ATG GAA GAA AAT GCA CCT AAG | 388 |
| Lys Ile Val Asp Asp Leu Val Ala Cys Met Glu Glu Asn Ala Pro Lys | |
| 85 90 95 | |
| AAT GTA AAA GAA TCA CTG AAG AAG CCA GAA ACT AGA AAC TTT ACT CCT | 436 |
| Asn Val Lys Glu Ser Leu Lys Lys Pro Glu Thr Arg Asn Phe Thr Pro | |
| 100 105 110 | |
| GAA GAA TTC TTT AGT ATT TTC AAT AGA TCC ATT GAT GCC TTC AAG GAC | 484 |
| Glu Glu Phe Phe Ser Ile Phe Asn Arg Ser Ile Asp Ala Phe Lys Asp | |
| 115 120 125 | |
| TTC ATG GTG GCA TCT GAC ACT AGT GAT TGT GTG CTC TCT TCA ACA TTA | 532 |
| Phe Met Val Ala Ser Asp Thr Ser Asp Cys Val Leu Ser Ser Thr Leu | |
| 130 135 140 | |
| GGT CCT GAG AAA GAT TCC AGA GTC AGT GTC ACA AAA CCA TTT ATG TTA | 580 |
| Gly Pro Glu Lys Asp Ser Arg Val Ser Val Thr Lys Pro Phe Met Leu | |
| 145 150 155 160 | |
| CCC CCT GTT GCA GCC AGT TCC CTT AGG AAT GAC AGC AGT AGC AGT AAT | 628 |
| Pro Pro Val Ala Ala Ser Ser Leu Arg Asn Asp Ser Ser Ser Ser Asn | |
| 165 170 175 | |
| AGG AAA GCC GCA AAG TCC CCT GAA GAC CCA GGC CTA CAA TGG ACA GCA | 676 |
| Arg Lys Ala Ala Lys Ser Pro Glu Asp Pro Gly Leu Gln Trp Thr Ala | |
| 180 185 190 | |
| ATG GCA CTG CCG GCT CTC ATT TCG CTT GTA ATT GGC TTT GCT TTT GGA | 724 |
| Met Ala Leu Pro Ala Leu Ile Ser Leu Val Ile Gly Phe Ala Phe Gly | |
| 195 200 205 | |
| GCC TTA TAC TGG AAG AAG AAA CAG TCA AGT CTT ACA AGG GCA GTT GAA | 772 |
| Ala Leu Tyr Trp Lys Lys Lys Gln Ser Ser Leu Thr Arg Ala Val Glu | |
| 210 215 220 | |
| AAT ATA CAG ATT AAT GAA GAG GAT AAT GAG ATA AGT ATG TTG CAA CAG | 820 |
| Asn Ile Gln Ile Asn Glu Glu Asp Asn Glu Ile Ser Met Leu Gln Gln | |
| 225 230 235 240 | |
| AAA GAG AGA GAG TTT CAA GAG GTG TAATT | 849 |
| Lys Glu Arg Glu Phe Gln Glu Val | |
| 245 | |

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 273 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

| |
|---|
| Met Lys Lys Thr Gln Thr Trp Ile Ile Thr Cys Ile Tyr Leu Gln Leu |
| -25 -20 -15 -10 |

Leu Leu Phe Asn Pro Leu Val Lys Thr Gln Glu Ile Cys Arg Asn Pro
 -5 1 5
 Val Thr Asp Asn Val Lys Asp Ile Thr Lys Leu Val Ala Asn Leu Pro
 10 15 20
 Asn Asp Tyr Met Ile Thr Leu Asn Tyr Val Ala Gly Met Asp Val Leu
 25 30 35
 Pro Ser His Cys Trp Leu Arg Asp Met Val Thr His Leu Ser Val Ser
 40 45 50 55
 Leu Thr Thr Leu Leu Asp Lys Phe Ser Asn Ile Ser Glu Gly Leu Ser
 60 65 70
 Asn Tyr Ser Ile Ile Asp Lys Leu Gly Lys Ile Val Asp Asp Leu Val
 75 80 85
 Ala Cys Met Glu Glu Asn Ala Pro Lys Asn Val Lys Glu Ser Leu Lys
 90 95 100
 Lys Pro Glu Thr Arg Asn Phe Thr Pro Glu Glu Phe Phe Ser Ile Phe
 105 110 115
 Asn Arg Ser Ile Asp Ala Phe Lys Asp Phe Met Val Ala Ser Asp Thr
 120 125 130 135
 Ser Asp Cys Val Leu Ser Ser Thr Leu Gly Pro Glu Lys Asp Ser Arg
 140 145 150
 Val Ser Val Thr Lys Pro Phe Met Leu Pro Pro Val Ala Ala Ser Ser
 155 160 165
 Leu Arg Asn Asp Ser Ser Ser Ser Asn Arg Lys Ala Ala Lys Ser Pro
 170 175 180
 Glu Asp Pro Gly Leu Gln Trp Thr Ala Met Ala Leu Pro Ala Leu Ile
 185 190 195
 Ser Leu Val Ile Gly Phe Ala Phe Gly Ala Leu Tyr Trp Lys Lys Lys
 200 205 210 215
 Gln Ser Ser Leu Thr Arg Ala Val Glu Asn Ile Gln Ile Asn Glu Glu
 220 225 230
 Asp Asn Glu Ile Ser Met Leu Gln Gln Lys Glu Arg Glu Phe Gln Glu
 235 240 245

Val

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3807 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: join(905..1018, 1914..1976, 2572..2742, 3152

..3307, 3513..3595)

(ix) FEATURE:

(A) NAME/KEY: mat_peptide

(B) LOCATION: join(965..1018, 1914..1976, 2572..2742, 3152
..3307, 3513..3595)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

| | |
|--|------|
| CACAAGTGAG TAGGGCGCGC CCGGGAGCTC CCAGGCTCTC CAGGAAAAAT CGCGCCCGGT | 60 |
| GCCCCGGGGA AGCCGGCGCT CCCTGGGACT TGCAGCTGGG GCGTGCAGGG CTGTGCCTGC | 120 |
| CGGGTGAGAT ACTACAAAGA TAAATCAGTT GCACAAGTTC TTGAAACTCT ACAGTGTAAT | 180 |
| AAGGAAAAAT AAGTCATGCA TAAAGCAAC TATAATACAT AATAGAAAAT GTTATTTTCA | 240 |
| AGCCGATGTG TAGGTTATGT GTGTTGAGA GAGAGAGAGA GAAGACAGAT TACTTTCTGC | 300 |
| TAGGGTTCAA GAATGCCTTC CTGTTGGCTA AGGAAATATT TTCCTTAAGT GGCTAAAAAG | 360 |
| CTGTGTTTCA AAATATTCTT TTGATGTCTC ACAAATTCAG TGAATTCTC TTAGGTCTAA | 420 |
| AAATATACAT CTCTCTCACT TTAACCTGGT GTGCTATTGT AGATTATTGG ATTAAAGCAC | 480 |
| TGCTCAGGGA TTATGCTGCT TCTTGCCAAG CAGTCTACAT TTAAAGTAGA AATAAGATGT | 540 |
| TTCTTTTGGT GCCATAAGGT ATACATTTTA TGCATTCTCT AGTTTTTAGA AGATACCCTA | 600 |
| AGGGCTAAGT CTTTAACATG CTGCTACAAG TTTATTCCTA ATTGCCATTG GGAAATTGGC | 660 |
| TGAAGAAAGT TTTTAACAAA AGTTAACAAAT ATTGTCATTG AGAGAATAAT TCAAAATGGA | 720 |
| TTTTAACTAA AAGCTTTTAA AAACCTTGGT GAGCATAGCT TGAATGCGTA ATATTTAATT | 780 |
| GCATTTAAGC CAATAACATA TATTAGACTG GTCTTTTTGT GCATCAAGGC ATTAGATGTT | 840 |
| AAAAGTTTGA ATGATTACAG ATCTTAACTG ATGATCACCA AGCAATTTTT CTGTTTTCAT | 900 |
| TTAG ACT TGG ATT CTC ACT TGC ATT TAT CTT CAG CTG CTC CTA TTT AAT | 949 |
| Thr Trp Ile Leu Thr Cys Ile Tyr Leu Gln Leu Leu Leu Phe Asn | |
| -20 -15 -10 | |
| CCT CTC GTC AAA ACT GAA GGG ATC TGC AGG AAT CGT GTG ACT AAT AAT | 997 |
| Pro Leu Val Lys Thr Glu Gly Ile Cys Arg Asn Arg Val Thr Asn Asn | |
| -5 1 5 10 | |
| GTA AAA GAC GTC ACT AAA TTG GTAAGTAAGG AATGCTTTAC CGTGCTGTGT | 1048 |
| Val Lys Asp Val Thr Lys Leu | |
| 15 | |
| AAAAAAGAGC TGTGGCTCTT TTTCTGTGC TTGTTGATAA AAGATTTAGA TTTTCTTGC | 1108 |
| CCCAAAGTAA TGTTTTCTTA AAGTGGGGAA AGTAATCACT GGGTTACAAT AAAGGGTTTA | 1168 |
| TAGAAAGCAG GTAGTGAGAT ATTTAGGGTC ATGGATAATT TGTTGGTAAA ACTGGCTAGT | 1228 |
| TGCACACCAC TGCTGTGACT GCTTCTTTGC TGGTCTTCTC CCCATCCTTC ATAGGCAGTG | 1288 |
| AAGGACCTTG GAGAGTTTCG TGTGTGCTGA TGGGCTTGCC CCAGCTTGTT CCCATAATC | 1348 |
| TCTCCAGTGG GTTTCCAGC ATGTTCTATT CCCCTTCA TGTCTTCTTA CTCTTCTTTA | 1408 |

| | |
|--|-------------|
| AAAAGCCTAA CGAAAGGAAA TCTGAAATGG CTATTCTCCC AATTCAATCA GCAGGAAGAC | 1468 |
| CCTGTACACAT GTCAGTGGGT GTTTGCTCCT TCAGGGAACA TAGAGAGGTG ATTCATTGCC | 1528 |
| CACATGTTGA AGGGACTCAT CTCCCTGGTT TGTCACATTG AACTCTTCCC TCAGCGAAAG | 1588 |
| CATTTGCATT GCTTCCCGAA TTCCAAGATC ACAGGTGGAA GCTGAAATTC AGATCATGTT | 1648 |
| TCCAAAACCTC AGTAGGTTAT ACCTAGCCAG GCATAACTGA ATTTGGAGTC TAAAAGATCT | 1708 |
| GTATTATCAC TTTTTTATTT TGAAGGATGC CTTTTGATTA CAGAGGGAAA TCAAGGATTA | 1768 |
| AAAATCAATA TACATGTAAA TATTGAAATT CATTGGTAAC TTTAAAAAGC ACAACAGTTT | 1828 |
| TGTGTGCTTT TCTCCAAAGC ACTACAAATA TGATTAATTG ATGTATAAGA ATTTTCTTAT | 1888 |
| GGAATTTTTT TTTTGTCTC TG TAG GTG GCA AAT CTT CCA AAA GAC TAC ATG | 1940 |
| Val Ala Asn Leu Pro Lys Asp Tyr Met | 20 25 |
| ATA ACC CTC AAA TAT GTC CCC GGG ATG GAT GTT TTG GTATGTAAAC | 1986 |
| Ile Thr Leu Lys Tyr Val Pro Gly Met Asp Val Leu | 30 35 |
| TACATTTCTG AGTTTCATTT TAGTAGCTCA TAGAAGAAAT GGGATCATTC ATATTGAGAT | 2046 |
| AGTACACTAG CTGCTATTTA GGAGCTTGCT TATTGTCAGG ATTTGAAGAA TTTATCTTTG | 2106 |
| GAATTTGACT TGCAGGCTTT TTTTCCCCC TCTTCCTGTT ACAAGAGTCC CTCCTCCTAT | 2166 |
| TACAATAGTC CCTCCTCCTC CTGTCACACT AGTCCCTTCT CTTCTGTGTA CAATAACCCC | 2226 |
| TGTCCTCCTA TTACAACATT TTAAGTAATG TAATATTAAT TTTAAAAATC TGGCCAGGCA | 2286 |
| CGGTGGTTCA TGCTTGTAAT CCCAGCACAT TGGGAAGCTG AGACGGGTGG ATCATTGAG | 2346 |
| GTCAGGAAGT TTGAGACAGC CTGGCCAACA TGGTGAAACT TCCTCTCTAC TAAAAATAAA | 2406 |
| AAAGTAGCCA GGCATGGTGG CAGGCACTTG TAATCTGAGC TACTCGAGAG GCTGAGGCAG | 2466 |
| GAGAATCACT TGAGTAACTA AAACGATAGC TTTGAAGAGT ACTCCGAGTT TTATGGCACT | 2526 |
| TACTTATTAA AATAGCTGTT TTGTCTCTTT TTTCATATCT TGCAG CCA AGT CAT | 2580 |
| Pro Ser His | 40 |
| TGT TGG ATA AGC GAG ATG GTA GTA CAA TTG TCA GAC AGC TTG ACT GAT | 2628 |
| Cys Trp Ile Ser Glu Met Val Val Gln Leu Ser Asp Ser Leu Thr Asp | 45 50 55 |
| CTT CTG GAC AAG TTT TCA AAT ATT TCT GAA GGC TTG AGT AAT TAT TCC | 2676 |
| Leu Leu Asp Lys Phe Ser Asn Ile Ser Glu Gly Leu Ser Asn Tyr Ser | 60 65 70 |
| ATC ATA GAC AAA CTT GTG AAT ATA GTG GAT GAC CTT GTG GAG TGC GTG | 2724 |
| Ile Ile Asp Lys Leu Val Asn Ile Val Asp Asp Leu Val Glu Cys Val | 75 80 85 90 |
| AAA GAA AAC TCA TCT AAG GTAACCTTGT GTTCATTGGG ATTATTTTTC | 2772 |
| Lys Glu Asn Ser Ser Lys | 95 |
| ATTACGCTTC TCTAAAAACC CATGCTTCTT GGTGCTGTTG GGGAAAATGA GGCACCTTTA | 2832 |

| | |
|--|------|
| TTTATGATAT TTTGATTGTA TAAACTTCAA ATTTAAAAAT CTTGTTTCAGA TGAGCAAAGA | 2892 |
| AAACAAGTAT TTGCAGTTAT ACTGCAATAC TGAAGTGCAC ATTCTTGTGT TCACTGCCCC | 2952 |
| AGATTCAACT TGTGATCCCA CTGGGATCAC TACCCTGCAT TACCAATCTG AATTACATAC | 3012 |
| GTTAAACAG CCATCTAAAA GTGCTAGTTG TAAGAGTCTA AATACTTGAA TCTTTGAGAG | 3072 |
| ACATATTTAT AGTCCATTAT CTTCACCTCA GTTAAGTCTG AAGACTATTT GAAAAATGTA | 3132 |
| ATCCTATTTT TTCTTCTAG GAT CTA AAA AAA TCA TTC AAG AGC CCA GAA CCC | 3184 |
| Asp Leu Lys Lys Ser Phe Lys Ser Pro Glu Pro | |
| 100 105 | |
| AGG CTC TTT ACT CCT GAA GAA TTC TTT AGA ATT TTT AAT AGA TCC ATT | 3232 |
| Arg Leu Phe Thr Pro Glu Glu Phe Phe Arg Ile Phe Asn Arg Ser Ile | |
| 110 115 120 | |
| GAT GCC TTC AAG GAC TTT GTA GTG GCA TCT GAA ACT AGT GAT TGT GTG | 3280 |
| Asp Ala Phe Lys Asp Phe Val Val Ala Ser Glu Thr Ser Asp Cys Val | |
| 125 130 135 | |
| GTT TCT TCA ACA TTA AGT CCT GAG AAA GGTAAGACAT GTAAGCATTT | 3327 |
| Val Ser Ser Thr Leu Ser Pro Glu Lys | |
| 140 145 | |
| CCAGTTCAAA TGTAACAAC AAACCTAAAT CTTCCCTATG TAGTAAGAAT CTACCTCTGT | 3387 |
| GTTAAGCTGT AGCAAGATAC ATGCATGTAC GTCTAATAAA AAAGCAGATA TCAATAGCAC | 3447 |
| AGAAGAACT CTATAACTCA TACAAATCAC CATATAACAC TGACACATTA TTGCTTTCTA | 3507 |
| TTTAG ATT CCA GAG TCA GTG TCA CAA AAC CAT TTA TGT TAC CCC CTG | 3554 |
| Ile Pro Glu Ser Val Ser Gln Asn His Leu Cys Tyr Pro Leu | |
| 150 155 160 | |
| TTG CAG CCA GCT CCC TTA GGA ATG ACA GCA GTA GCA GTA ATAGTAAGTA | 3603 |
| Leu Gln Pro Ala Pro Leu Gly Met Thr Ala Val Ala Val | |
| 165 170 175 | |
| CATATATCTG ATTTAATGCA TGCATGGCTC CAATTAGCAC CTATAGGAGT ATTGCATGGG | 3663 |
| CTTTCAAGGA AACTTCTACA TTTATTATTA TTGATACTGT TCTGTACTG TTATTCCTTT | 3723 |
| TATGGTCTTC TTGAGACTTA AGTTTGTAGA ATTAAATTTT CCTAGAGCTG GAGATAATGT | 3783 |
| TTAGAGAATT AGGCCAATAA ATTT | 3807 |

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 195 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

| |
|---|
| Thr Trp Ile Leu Thr Cys Ile Tyr Leu Gln Leu Leu Leu Phe Asn Pro |
| -20 -15 -10 -5 |

| | |
|---|-----|
| ATC TGC AGG AAT CGT GTG ACT AAT AAT GTA AAA GAC GTC ACT AAA TTG Ile Cys Arg Asn Arg Val Thr Asn Asn Val Lys Asp Val Thr Lys Leu 5 10 15 | 145 |
| GTG GCA AAT CTT CCA AAA GAC TAC ATG ATA ACC CTC AAA TAT GTC CCC Val Ala Asn Leu Pro Lys Asp Tyr Met Ile Thr Leu Lys Tyr Val Pro 20 25 30 | 193 |
| GGG ATG GAT GTT TTG CCA AGT CAT TGT TGG ATA AGC GAG ATG GTA GTA Gly Met Asp Val Leu Pro Ser His Cys Trp Ile Ser Glu Met Val Val 35 40 45 50 | 241 |
| CAA TTG TCA GAC AGC TTG ACT GAT CTT CTG GAC AAG TTT TCA AAT ATT Gln Leu Ser Asp Ser Leu Thr Asp Leu Leu Asp Lys Phe Ser Asn Ile 55 60 65 | 289 |
| TCT GAA GGC TTG AGT AAT TAT TCC ATC ATA GAC AAA CTT GTG AAT ATA Ser Glu Gly Leu Ser Asn Tyr Ser Ile Ile Asp Lys Leu Val Asn Ile 70 75 80 | 337 |
| GTG GAT GAC CTT GTG GAG TGC GTG AAA GAA AAC TCA TCT AAG GAT CTA Val Asp Asp Leu Val Glu Cys Val Lys Glu Asn Ser Ser Lys Asp Leu 85 90 95 | 385 |
| AAA AAA TCA TTC AAG AGC CCA GAA CCC AGG CTC TTT ACT CCT GAA GAA Lys Lys Ser Phe Lys Ser Pro Glu Pro Arg Leu Phe Thr Pro Glu Glu 100 105 110 | 433 |
| TTC TTT AGA ATT TTT AAT AGA TCC ATT GAT GCC TTC AAG GAC TTT GTA Phe Phe Arg Ile Phe Asn Arg Ser Ile Asp Ala Phe Lys Asp Phe Val 115 120 125 130 | 481 |
| GTG GCA TCT GAA ACT AGT GAT TGT GTG GTT TCT TCA ACA TTA AGT CCT Val Ala Ser Glu Thr Ser Asp Cys Val Val Ser Ser Thr Leu Ser Pro 135 140 145 | 529 |
| GAG AAA GAT TCC AGA GTC AGT GTC ACA AAA CCA TTT ATG TTA CCC CCT Glu Lys Asp Ser Arg Val Ser Val Thr Lys Pro Phe Met Leu Pro Pro 150 155 160 | 577 |
| GTT GCA GCC AGC TCC CTT AGG AAT GAC AGC AGT AGC AGT AAT AGT AAG Val Ala Ala Ser Ser Leu Arg Asn Asp Ser Ser Ser Ser Asn Ser Lys 165 170 175 | 625 |
| TAC ATA TAT CTG ATT TAATGCATGC ATGGCTCCAA TTAGCACCTA TAGGAGTATT Tyr Ile Tyr Leu Ile 180 | 680 |
| GCATGGGCTT TCAAGGAAAC TTCTACATTT ATTATTATTG ATACTGTTCT GTTACTGTTA | 740 |
| TTCTTTTAT GGTCTTCTTG AGACTTAAGT TTGTAGAATT AAATTTCCCT AGAGCTGGAG | 800 |
| ATAATGTTTA GAGAATTAGG | 820 |

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 208 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

Met Lys Lys Thr Gln Thr Trp Ile Leu Thr Cys Ile Tyr Leu Gln Leu
-25 -20 -15 -10
Leu Leu Phe Asn Pro Leu Val Lys Thr Glu Gly Ile Cys Arg Asn Arg
-5 1 5
Val Thr Asn Asn Val Lys Asp Val Thr Lys Leu Val Ala Asn Leu Pro
10 15 20
Lys Asp Tyr Met Ile Thr Leu Lys Tyr Val Pro Gly Met Asp Val Leu
25 30 35
Pro Ser His Cys Trp Ile Ser Glu Met Val Val Gln Leu Ser Asp Ser
40 45 50 55
Leu Thr Asp Leu Leu Asp Lys Phe Ser Asn Ile Ser Glu Gly Leu Ser
60 65 70
Asn Tyr Ser Ile Ile Asp Lys Leu Val Asn Ile Val Asp Asp Leu Val
75 80 85
Glu Cys Val Lys Glu Asn Ser Ser Lys Asp Leu Lys Lys Ser Phe Lys
90 95 100
Ser Pro Glu Pro Arg Leu Phe Thr Pro Glu Glu Phe Phe Arg Ile Phe
105 110 115
Asn Arg Ser Ile Asp Ala Phe Lys Asp Phe Val Val Ala Ser Glu Thr
120 125 130 135
Ser Asp Cys Val Val Ser Ser Thr Leu Ser Pro Glu Lys Asp Ser Arg
140 145 150
Val Ser Val Thr Lys Pro Phe Met Leu Pro Pro Val Ala Ala Ser Ser
155 160 165
Leu Arg Asn Asp Ser Ser Ser Ser Asn Ser Lys Tyr Ile Tyr Leu Ile
170 175 180

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5864 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: join(565..579, 1684..1797, 2693..2755, 3351..3521, 3932..4088, 4314..4397, 4778..4887, 5208..5275, 5677..5713)

(ix) FEATURE:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION: join(1744..1797, 2693..2755, 3351..3521, 3932..4088, 4314..4397, 4778..4887, 5208..5275, 5677..5713)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

| | |
|--|------|
| GAGCTCCGAG CCCTCTCTGG CGCGCGAGGT ATTTCTGTCTG TNCCCGGGGG TGCCAGGTGA | 60 |
| GCCCCAGCGG ATCCGGGAGG GTAAGCTGGG ACTCCTCGCG AGCAGTAGCT GCAGGGTACC | 120 |
| AAGCTTCGCC CTCTGCGTCC CCGCGCCTTC GCGGTCTCCC GCCAGTGCAG GTCCGGGGCC | 180 |
| CCCAGGCGAG CGGACAAGGT TGGCCTAATC TGCCAAACTT CTGGGGCATT TACCGTGCTC | 240 |
| TGGCCGCCCT CCCGATTCTT CCCTCCGCGC CTTGCGCTGC TTCTCGCCTA CCCCAGGGCTC | 300 |
| CGGAAGGGAA GGAGGCGTGT CCGGAGCAGG CGGGCGGGAA CTGTATAAAA GCGCCGCGCG | 360 |
| CTCAGCAGCC GGCTTCGCTC GCCGCCTCGC GCCGAGACTA GAAGCGCTGC GGGAAGCAGG | 420 |
| GACAGTGGAG AGGGCGCTGC GCTCGGGCTA CCCAATGCGT GGAATATCTG CCGCCGCTGT | 480 |
| TCGTGCAATA TGCTGGAGCT CCAGAACAGC TAAACGGAGT CGCCACACCA CTGTTTGTGC | 540 |
| TGGATCGCAC CGCTGCCTTT CCTT ATG AAG AAG ACA CAA GTGAGTAGGG | 589 |
| Met Lys Lys Thr Gln | |
| -25 | |
| CGCGCCCGGG AGCTCCCAGG CTCTCCAGGA AAAATCGCGC CCGGTGCCCC GGGGAAGCCG | 649 |
| GCGCTCCCTG GGACTTGCA GCTGGGGCGTG CAGGGCTGTG CCTGCCGGGT GAGACAAGAG | 709 |
| GATGCGGGGG AGGCCGGCGT GGTGTGTGAT CCCGAGCCGA GCCGNNTGAG CCAGGGAGAA | 769 |
| AAGGAGTGGG AGTACTGAGA GGGAGCCAGT GTCAAGTTTG GAGCCTCAGC AGTTAAGTTT | 829 |
| TGAGCTGTCA GTCGGAAACC GTAATTCCCG TCTGGTGGAA AGATTGGCTT TTNGNCCACG | 889 |
| GAATGTAAGT TATCACAGAT ACTACAAAGA TAAATCAGTT GCACAAGTTC TTGAAACTCT | 949 |
| ACAGTGTAAT AAGGAAAAAT AAGTCATGCA TAAAAGCAAC TATAATACAT AATAGAAAAT | 1009 |
| GTTATTTTCA AGCCGATGTG TAGGTTATGT GTGTTCGAGA GAGAGAGAGA GAAGACAGAT | 1069 |
| TACTTTCTGC TAGGGTTCAA GAATGCCTTC CTGTTGGCTA AGGAAATATT TTCCTTAAGT | 1129 |
| GGCTAAAAAG CTGTGTTTCA AAATATTCTT TTGATGTCTC ACAAATTCAG TGGAATTCTC | 1189 |
| TTAGGTCTAA AAATATACAT CTCTCTCACT TTAAGTTGGT GTGCTATTGT AGATTATTGG | 1249 |
| ATTAAAGCAC TGCTCAGGGA TTATGCTGCT TCTTGCCAAG CAGTCTACAT TTAAAGTAGA | 1309 |
| AATAAGATGT TTCTTTTGGT GCCATAAGGT ATACATTTTA TGCATTCTCT AGTTTTTAGA | 1369 |
| AGATACCCTA AGGGCTAAGT CTTTAACATG CTGCTACAAG TTTATTCCTA ATTGCCATTG | 1429 |
| GGAAATTGGC TGAAGAAAGT TTTTAACAAA AGTTAACAAT ATTGTCATTG AGAGAATAAT | 1489 |
| TCAAAATGGA TTTTAACTAA AAGCTTTTAA AACTTTGGT GAGCATAGCT TGAATGCGTA | 1549 |
| ATATTTAATT GCATTTAAGC CAATAACATA TATTAGACTG GTCTTTTGT GCATCAAGGC | 1609 |
| ATTAGATGTT AAAAGTTTGA ATGATTACAG ATCTTAACTG ATGATCACCA AGCAATTTTT | 1669 |
| CTGTTTTTCAT TTAG ACT TGG ATT CTC ACT TGC ATT TAT CTT CAG CTG CTC | 1719 |
| Thr Trp Ile Leu Thr Cys Ile Tyr Leu Gln Leu Leu | |
| -20 -15 -10 | |

| | |
|--|------|
| CTA TTT AAT CCT CTC GTC AAA ACT GAA GGG ATC TGC AGG AAT CGT GTG Leu Phe Asn Pro Leu Val Lys Thr Glu Gly Ile Cys Arg Asn Arg Val | 1767 |
| -5 1 5 | |
| ACT AAT AAT GTA AAA GAC GTC ACT AAA TTG GTAAGTAAGG AATGCTTTAC Thr Asn Asn Val Lys Asp Val Thr Lys Leu | 1817 |
| 10 15 | |
| CGTGCTGTGT AAAAAAGAGC TGTGGCTCTT TTTCCTGTGC TTGTTGATAA AAGATTTAGA | 1877 |
| TTTTTCTTGC CCCAAAGTAA TGTTTTCTTA AAGTGGGGAA AGTAATCACT GGGTTACAAT | 1937 |
| AAAGGGTTTA TAGAAAGCAG GTAGTGAGAT ATTTAGGGTC ATGGATAATT TGTTGGTAA | 1997 |
| ACTGGCTAGT TGCACACCAC TGCTGTGACT GCTTCTTTGC TGGTCTTCTC CCCATCCTTC | 2057 |
| ATAGGCAGTG AAGGACCTTG GAGAGTTCGC TGTGTGCTGA TGGGCTTGCC CCAGCTTGTT | 2117 |
| CCCCATAATC TCTCCAGTGG GTTTCCTCAGC ATGTTCTATT CCCCTTCACA TGTCTTCCTA | 2177 |
| CTCTTCTTTA AAAAGCCTAA CGAAAGGAAA TCTGAAATGG CTATTCTCCC AATTCAATCA | 2237 |
| GCAGGAAGAC CCTGTCACAT GTCAGTGGGT GTTTGCTCCT TCAGGGAACA TAGAGAGGTG | 2297 |
| ATTCATTGCC CACATGTTGA AGGGACTCAT CTCCCTGGTT TGTCACATTG AACTCTTCCC | 2357 |
| TCAGCGAAAG CATTTGCATT GCTTCCCGAA TTCCAAGATC ACAGGTGGAA GCTGAAATTC | 2417 |
| AGATCATGTT TCCAAACTC AGTAGGTTAT ACCTAGCCAG GCATAACTGA ATTTGGAGTC | 2477 |
| TAAAAGATCT GTATTATCAC TTTTTATTT TGAAGGATGC CTTTTGATTA CAGAGGGAAA | 2537 |
| TCAAGGATTA AAAATCAATA TACATGTAAA TATTGAAATT CATTGGTAAC TTTAAAAAGC | 2597 |
| ACAACAGTTT TGTGTGCTTT TCTCCAAAGC ACTACAAATA TGATTAATTG ATGTATAAGA | 2657 |
| ATTTTCTTAT GGAATTTTTT TTTTGTCTC TGTAG GTG GCA AAT CTT CCA AAA Val Ala Asn Leu Pro Lys | 2710 |
| 20 | |
| GAC TAC ATG ATA ACC CTC AAA TAT GTC CCC GGG ATG GAT GTT TTG Asp Tyr Met Ile Thr Leu Lys Tyr Val Pro Gly Met Asp Val Leu | 2755 |
| 25 30 35 | |
| GTATGTAAAC TACATTTCTG AGTTTCATTT TAGTAGCTCA TAGAAGAAAT GGGATCATTC | 2815 |
| ATATTGAGAT AGTACACTAG CTGCTATTTA GGAGCTTGCT TATTGTCAGG ATTTGAAGAA | 2875 |
| TTTATCTTTG GAATTTGACT TGCAGGCTTT TTTTCCCCC TCTTCCTGTT ACAAGAGTCC | 2935 |
| CTCCTCCTAT TACAATAGTC CCTCCTCCTC CTGTCACACT AGTCCCTTCT CTCCTGTTA | 2995 |
| CAATAACCCC TGTCCTCCTA TTACAACATT TTAAGTAATG TAATATTAAT TTTAAAAATC | 3055 |
| TGGCCAGGCA CGGTGGTTCA TGCTTGTAAT CCCAGCACAT TGGGAAGCTG AGACGGGTGG | 3115 |
| ATCATTTGAG GTCAGGAAGT TTGAGACAGC CTGGCCAACA TGGTGAAACT TCCTCTCTAC | 3175 |
| TAAAAATAAA AAAGTAGCCA GGCATGGTGG CAGGCACTTG TAATCTGAGC TACTCGAGAG | 3235 |
| GCTGAGGCAG GAGAATCACT TGAGTAACTA AAACGATAGC TTTGAAGAGT ACTCCGAGTT | 3295 |
| TTATGGCACT TACTTATTAA AATAGCTGTT TTGTCTCTTT TTTCATATCT TGCAG CCA | 3353 |

Pro
40

| | |
|---|------|
| AGT CAT TGT TGG ATA AGC GAG ATG GTA GTA CAA TTG TCA GAC AGC TTG Ser His Cys Trp Ile Ser Glu Met Val Val Gln Leu Ser Asp Ser Leu 45 50 55 | 3401 |
| ACT GAT CTT CTG GAC AAG TTT TCA AAT ATT TCT GAA GGC TTG AGT AAT Thr Asp Leu Leu Asp Lys Phe Ser Asn Ile Ser Glu Gly Leu Ser Asn 60 65 70 | 3449 |
| TAT TCC ATC ATA GAC AAA CTT GTG AAT ATA GTG GAT GAC CTT GTG GAG Tyr Ser Ile Ile Asp Lys Leu Val Asn Ile Val Asp Asp Leu Val Glu 75 80 85 | 3497 |
| TGC GTG AAA GAA AAC TCA TCT AAG GTAACCTTTGT GTTCATTGGG ATTATTTTTTC Cys Val Lys Glu Asn Ser Ser Lys 90 95 | 3551 |
| ATTACGCTTC TCTAAAAACC CATGCTTCTT GGTGCTGTTG GGGAAAATGA GGCACCTTTA | 3611 |
| TTTATGATAT TTTGATTGTA TAAACTTCAA ATTTAAAAAT CTTGTTTCAGA TGAGCAAAGA | 3671 |
| AAACAAGTAT TTGCAGTTAT ACTGCAATAC TGAAGTGCAC TATTCTTGTG TTCACTGCCC | 3731 |
| CAGATTCAAC TTGTGATCCC ACTGGGATCA CTACCCTGCA TTACCAATCT GAATTACATA | 3791 |
| CGTTAAAACA GCCATCTAAA AGTGCTAGTT GTAAGAGTCT AAATACTTGA ATCTTTGAGA | 3851 |
| GACATATTTA TAGTCCATTA TCTTCACCTC AGTTAAGTCT GAAGACTATT TGAAAAATGT | 3911 |
| AATCCTATTT TTTCTTCTAG GAT CTA AAA AAA TCA TTC AAG AGC CCA GAA Asp Leu Lys Lys Ser Phe Lys Ser Pro Glu 100 105 | 3961 |
| CCC AGG CTC TTT ACT CCT GAA GAA TTC TTT AGA ATT TTT AAT AGA TCC Pro Arg Leu Phe Thr Pro Glu Glu Phe Phe Arg Ile Phe Asn Arg Ser 110 115 120 | 4009 |
| ATT GAT GCC TTC AAG GAC TTT GTA GTG GCA TCT GAA ACT AGT GAT TGT Ile Asp Ala Phe Lys Asp Phe Val Val Ala Ser Glu Thr Ser Asp Cys 125 130 135 | 4057 |
| GTG GTT TCT TCA ACA TTA AGT CCT GAG AAA G GTAAGACATG TAAGCATTTTC Val Val Ser Ser Thr Leu Ser Pro Glu Lys 140 145 | 4108 |
| CAGTTCAAAT GTAAACAACA AACTTAAATC TTCCCTATGT AGTAAGAATC TACCTCTGTG | 4168 |
| TTAAGCTGTA GCAAGATACA TGCATGTACG TCTAAAAAAA AGCAGATATC AATAGCACAG | 4228 |
| AAGAACTAA TGATTGTAGA TTTGTGGGCT CTATAACTCA TACAAATCAC CATATAACAC | 4288 |
| TGACACATTA TTGCTTTCTA TTTAG AT TCC AGA GTC AGT GTC ACA AAA CCA Asp Ser Arg Val Ser Val Thr Lys Pro 150 155 | 4339 |
| TTT ATG TTA CCC CCT GTT GCA GCC AGC TCC CTT AGG AAT GAC AGC AGT Phe Met Leu Pro Pro Val Ala Ala Ser Ser Leu Arg Asn Asp Ser Ser 160 165 170 | 4387 |

| | |
|---|------|
| AGC AGT AAT A GTAAGTACAT ATATCTGATT TAATGCATGC ATGGCTCCAA Ser Ser Asn 175 | 4437 |
| TTAGCACCTA TAGGAGTATT GCATGGGCTT TCAAGGAAAC TTCTACATTT ATTATTATTG | 4497 |
| ATACTGTTCT GTTACTGTTA TTCCTTTTAT GGTCTTCTTG AGACTTAAGT TTGTAGAATT | 4557 |
| AAATTTCCCT AGAGCTGGAG ATAATGTTTA GAGAATTAGG CCAATAAATT TTCTGCTGAG | 4617 |
| GTTATTTTAA ATAAGACATA AAATTAATTT TAGAAATATG ATTTATGCCT TTTGTTGAAT | 4677 |
| CATTAACATA TATACAGAAA CAGTTAAAAC AACCACAGCA TAAGAGAAAA ACTTCTAGAA | 4737 |
| TGGATATGCT GTATTCATCA GTGTGTTCTT TAAATTATAG GG AAG GCC AAA AAT Arg Lys Ala Lys Asn 180 | 4791 |
| CCC CCT GGA GAC TCC AGC CTA CAC TGG CCA GCC ATG GCA TTG CCA GCA Pro Pro Gly Asp Ser Ser Leu His Trp Pro Ala Met Ala Leu Pro Ala 185 190 195 | 4839 |
| TTG TTT TCT CTT ATA ATT GGC TTT GCT TTT GGA GCC TTA TAC TGG AAG Leu Phe Ser Leu Ile Ile Gly Phe Ala Phe Gly Ala Leu Tyr Trp Lys 200 205 210 | 4887 |
| GTAAGTGGTA CCATTCCTTT TTTTAAAAAT ATGCTATGTT TACATAAATT ATCATCTTTT | 4947 |
| TTTCCTCAAG AAATGATCCT TTAAGAAAAC AGTGAATCTA CCTTAGCTTA TACTAAACAA | 5007 |
| AATTTAAATT TTATAAAGTT TCCTGTTTCT CATTATGTCT GGAGACAATC CCTCTAGCTG | 5067 |
| ATAATTCACG CTTAAGAATT AGGAACTAAA ACTGTTATTG GAGTTATTGC CATAAAAGAT | 5127 |
| AAAAGTGGAG TCCACTTACC TCTTAAATAT TAGACCATTC ATTGATTATT TTACAGTATA | 5187 |
| TGTCTTTCTT CTTTTTCCAG AAG AGA CAG CCA AGT CTT ACA AGG GCA GTT Lys Arg Gln Pro Ser Leu Thr Arg Ala Val 215 220 | 5237 |
| GAA AAT ATA CAA ATT AAT GAA GAG GAT AAT GAG ATA AG GTATTTTGTT Glu Asn Ile Gln Ile Asn Glu Glu Asp Asn Glu Ile Ser 225 230 235 | 5285 |
| TTGCTAAATG TGTGCCCAAT CAAGCATGAC ATTGCCATTT CACACACTGT GTACCTGCCC | 5345 |
| ATAATGTCTT TAAGAAGTCC TTTACTCATG ACAGTAGCTC CTAACCAAGTG AGTCCCAACT | 5405 |
| CTATCCATGT TTCTGATGTC TCACTCTCTC TTCGTATGTG TATATGCATA TACAGAGAAA | 5465 |
| GAAATGTTTT AACTACTTGG AAAGACTACC TTAAGACAAA TGAAGTCTTC CCTCTTCCCT | 5525 |
| ATAGTAATAA GAAGGTAGGC TCCCCATTC AATTTTGCAA TCTTCTGCTA CTATATTTAC | 5585 |
| AGAAAAGCTG CCTTTTACAA TGCCGAGATC ATGGTGTACC TCAGAATCTC TGACCAAGAG | 5645 |
| CAAATAAGCA TTTTTTCTTA TTGTTTTTCA G T ATG TTG CAA GAG AAA GAG AGA Met Leu Gln Glu Lys Glu Arg 240 | 5698 |
| GAG TTT CAA GAA GTG TAATTGTGGC TTGTATCAAC ACTGTTACTT TCGTACATTG Glu Phe Gln Glu Val 245 | 5753 |

GTAAGTTTTT TTCTTCTTTC CTTTTTTTTT CTTTTTTTTA TTATACTTTA AGTTCTAGGG 5813
TACATGTGCA CAATGTGCAG GTTTGTTACG TATGTTTACA TGTGCCATGT T 5864

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 273 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

Met Lys Lys Thr Gln Thr Trp Ile Leu Thr Cys Ile Tyr Leu Gln Leu
-25 -20 -15 -10
Leu Leu Phe Asn Pro Leu Val Lys Thr Glu Gly Ile Cys Arg Asn Arg
-5 1 5
Val Thr Asn Asn Val Lys Asp Val Thr Lys Leu Val Ala Asn Leu Pro
10 15 20
Lys Asp Tyr Met Ile Thr Leu Lys Tyr Val Pro Gly Met Asp Val Leu
25 30 35
Pro Ser His Cys Trp Ile Ser Glu Met Val Val Gln Leu Ser Asp Ser
40 45 50 55
Leu Thr Asp Leu Leu Asp Lys Phe Ser Asn Ile Ser Glu Gly Leu Ser
60 65 70
Asn Tyr Ser Ile Ile Asp Lys Leu Val Asn Ile Val Asp Asp Leu Val
75 80 85
Glu Cys Val Lys Glu Asn Ser Ser Lys Asp Leu Lys Lys Ser Phe Lys
90 95 100
Ser Pro Glu Pro Arg Leu Phe Thr Pro Glu Glu Phe Phe Arg Ile Phe
105 110 115
Asn Arg Ser Ile Asp Ala Phe Lys Asp Phe Val Val Ala Ser Glu Thr
120 125 130 135
Ser Asp Cys Val Val Ser Ser Thr Leu Ser Pro Glu Lys Asp Ser Arg
140 145 150
Val Ser Val Thr Lys Pro Phe Met Leu Pro Pro Val Ala Ala Ser Ser
155 160 165
Leu Arg Asn Asp Ser Ser Ser Ser Asn Arg Lys Ala Lys Asn Pro Pro
170 175 180
Gly Asp Ser Ser Leu His Trp Pro Ala Met Ala Leu Pro Ala Leu Phe
185 190 195
Ser Leu Ile Ile Gly Phe Ala Phe Gly Ala Leu Tyr Trp Lys Lys Arg
200 205 210 215
Gln Pro Ser Leu Thr Arg Ala Val Glu Asn Ile Gln Ile Asn Glu Glu
220 225 230

Asp Asn Glu Ile Ser Met Leu Gln Glu Lys Glu Arg Glu Phe Gln Glu
 235 240 245

Val

(2) INFORMATION FOR SEQ ID NO:49:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 273 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Lys | Lys | Thr | Gln | Thr | Trp | Ile | Leu | Thr | Cys | Ile | Tyr | Leu | Gln | Leu | 1 | 5 | 10 | 15 |
| Leu | Leu | Phe | Asn | Pro | Leu | Val | Lys | Thr | Glu | Gly | Ile | Cys | Arg | Asn | Arg | 20 | 25 | 30 | |
| Val | Thr | Asn | Asn | Val | Lys | Asp | Val | Thr | Lys | Leu | Val | Ala | Asn | Leu | Pro | 35 | 40 | 45 | |
| Lys | Asp | Tyr | Met | Ile | Thr | Leu | Lys | Tyr | Val | Pro | Gly | Met | Asp | Val | Leu | 50 | 55 | 60 | |
| Pro | Ser | His | Cys | Trp | Ile | Ser | Glu | Met | Val | Val | Gln | Leu | Ser | Asp | Ser | 65 | 70 | 75 | 80 |
| Leu | Thr | Asp | Leu | Leu | Asp | Lys | Phe | Ser | Asn | Ile | Ser | Glu | Gly | Leu | Ser | 85 | 90 | 95 | |
| Asn | Tyr | Ser | Ile | Ile | Asp | Lys | Leu | Val | Asn | Ile | Val | Asp | Asp | Leu | Val | 100 | 105 | 110 | |
| Glu | Cys | Val | Lys | Glu | Asn | Ser | Ser | Lys | Asp | Leu | Lys | Lys | Ser | Phe | Lys | 115 | 120 | 125 | |
| Ser | Pro | Glu | Pro | Arg | Leu | Phe | Thr | Pro | Glu | Glu | Phe | Phe | Arg | Ile | Phe | 130 | 135 | 140 | |
| Asn | Arg | Ser | Ile | Asp | Ala | Phe | Lys | Asp | Phe | Val | Val | Ala | Ser | Glu | Thr | 145 | 150 | 155 | 160 |
| Ser | Asp | Cys | Val | Val | Ser | Ser | Thr | Leu | Ser | Pro | Glu | Lys | Asp | Ser | Arg | 165 | 170 | 175 | |
| Val | Ser | Val | Thr | Lys | Pro | Phe | Met | Leu | Pro | Pro | Val | Ala | Ala | Ser | Ser | 180 | 185 | 190 | |
| Leu | Arg | Asn | Asp | Ser | Ser | Ser | Ser | Asn | Arg | Lys | Ala | Lys | Asn | Pro | Pro | 195 | 200 | 205 | |
| Gly | Asp | Ser | Ser | Leu | His | Trp | Ala | Ala | Met | Ala | Leu | Pro | Ala | Leu | Phe | 210 | 215 | 220 | |
| Ser | Leu | Ile | Ile | Gly | Phe | Ala | Phe | Gly | Ala | Leu | Tyr | Trp | Lys | Lys | Arg | 225 | 230 | 235 | 240 |

Gln Pro Ser Leu Thr Arg Ala Val Glu Asn Ile Gln Ile Asn Glu Glu
245 250 255
Asp Asn Glu Ile Ser Met Leu Gln Glu Lys Glu Arg Glu Phe Gln Glu
260 265 270
Val

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 273 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

Met Lys Lys Thr Gln Thr Trp Ile Leu Thr Cys Ile Tyr Leu Gln Leu
1 5 10 15
Leu Leu Phe Asn Pro Leu Val Lys Thr Glu Gly Ile Cys Arg Asn Arg
20 25 30
Val Thr Asn Asn Val Lys Asp Val Thr Lys Leu Val Ala Asn Leu Pro
35 40 45
Lys Asp Tyr Met Ile Thr Leu Lys Tyr Val Pro Gly Met Asp Val Leu
50 55 60
Pro Ser His Cys Trp Ile Ser Glu Met Val Val Gln Leu Ser Asp Ser
65 70 75 80
Leu Thr Asp Leu Leu Asp Lys Phe Ser Asn Ile Ser Glu Gly Leu Ser
85 90 95
Asn Tyr Ser Ile Ile Asp Lys Leu Val Asn Ile Val Asp Asp Leu Val
100 105 110
Glu Cys Val Lys Glu Asn Ser Ser Lys Asp Leu Lys Lys Ser Phe Lys
115 120 125
Ser Pro Glu Pro Arg Leu Phe Thr Pro Glu Glu Phe Phe Arg Ile Phe
130 135 140
Asn Arg Ser Ile Asp Ala Phe Lys Asp Phe Ala Val Ala Ser Glu Thr
145 150 155 160
Ser Asp Cys Val Val Ser Ser Thr Leu Ser Pro Glu Lys Asp Ser Arg
165 170 175
Val Ser Val Thr Lys Pro Phe Met Leu Pro Pro Val Ala Ala Ser Ser
180 185 190
Leu Arg Asn Asp Ser Ser Ser Ser Asn Arg Lys Ala Lys Asn Pro Thr
195 200 205
Gly Asp Ser Ser Leu His Trp Ala Ala Met Ala Leu Pro Ala Phe Phe
210 215 220

Ser Leu Ile Ile Gly Phe Ala Phe Gly Ala Leu Tyr Trp Lys Lys Arg
 225 230 235 240
 Gln Pro Ser Leu Thr Arg Ala Val Glu Asn Ile Gln Ile Asn Glu Glu
 245 250 255
 Asp Asn Glu Ile Ser Met Leu Gln Glu Lys Glu Arg Glu Phe Gln Glu
 260 265 270
 Val

(2) INFORMATION FOR SEQ ID NO:51:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 274 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

Met Lys Lys Thr Gln Thr Trp Ile Ile Thr Cys Ile Tyr Leu Gln Leu
 1 5 10 15
 Leu Leu Phe Asn Pro Leu Val Lys Thr Lys Gly Ile Cys Gly Lys Arg
 20 25 30
 Val Thr Asp Asp Val Lys Asp Val Thr Lys Leu Val Ala Asn Leu Pro
 35 40 45
 Lys Asp Tyr Lys Ile Ala Leu Lys Tyr Val Pro Gly Met Asp Val Leu
 50 55 60
 Pro Ser His Cys Trp Ile Ser Val Met Val Glu Gln Leu Ser Val Ser
 65 70 75 80
 Leu Thr Asp Leu Leu Asp Lys Phe Ser Asn Ile Ser Glu Gly Leu Ser
 85 90 95
 Asn Tyr Ser Ile Ile Asp Lys Leu Val Lys Ile Val Asp Asp Leu Val
 100 105 110
 Glu Cys Thr Glu Gly Tyr Ser Phe Glu Asn Val Lys Lys Ala Pro Lys
 115 120 125
 Ser Pro Glu Leu Arg Leu Phe Thr Pro Glu Glu Phe Phe Arg Ile Phe
 130 135 140
 Asn Arg Ser Ile Asp Ala Phe Lys Asp Leu Glu Thr Val Ala Ser Lys
 145 150 155 160
 Ser Ser Glu Cys Val Val Ser Ser Thr Leu Ser Pro Asp Lys Asp Ser
 165 170 175
 Arg Val Ser Val Thr Lys Pro Phe Met Leu Pro Pro Val Ala Ala Ser
 180 185 190
 Ser Leu Arg Asn Asp Ser Ser Ser Ser Asn Arg Lys Ala Ser Asn Ser
 195 200 205

Ile Gly Asp Ser Asn Leu Gln Trp Ala Ala Met Ala Leu Pro Ala Phe
 210 215 220
 Phe Ser Leu Val Ile Gly Phe Ala Phe Gly Ala Leu Tyr Trp Lys Lys
 225 230 235 240
 Lys Gln Pro Asn Leu Thr Arg Thr Val Glu Asn Ile Gln Ile Asn Glu
 245 250 255
 Glu Asp Asn Glu Ile Ser Met Leu Gln Glu Lys Glu Arg Glu Phe Gln
 260 265 270
 Glu Val

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 271 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

Met Lys Thr Gln Thr Trp Ile Val Thr Cys Ile Tyr Leu Gln Leu Leu
 1 5 10 15
 Phe Asn Pro Leu Val Lys Thr Lys Gly Leu Cys Arg Asn Arg Val Thr
 20 25 30
 Asp Asp Val Lys Asp Val Thr Lys Leu Val Ala Asn Leu Pro Lys Asp
 35 40 45
 Tyr Lys Ile Ala Leu Lys Tyr Val Pro Gly Met Asp Val Leu Pro Ser
 50 55 60
 His Cys Trp Ile Ser Val Met Val Glu Gln Leu Ser Val Ser Leu Thr
 65 70 75 80
 Asp Leu Leu Asp Lys Phe Ser Asn Ile Ser Glu Gly Leu Ser Asn Tyr
 85 90 95
 Ser Ile Ile Asp Lys Leu Val Lys Ile Val Asp Asp Leu Val Glu Cys
 100 105 110
 Val Glu Gly His Ser Ser Glu Asn Val Lys Lys Ser Ser Lys Ser Pro
 115 120 125
 Glu Pro Arg Leu Phe Thr Pro Glu Glu Phe Phe Arg Ile Phe Asn Arg
 130 135 140
 Ser Ile Asp Ala Phe Lys Asp Leu Glu Met Val Ala Ser Lys Thr Ser
 145 150 155 160
 Glu Cys Val Val Ser Ser Thr Leu Ser Pro Glu Lys Asp Ser Arg Val
 165 170 175
 Ser Val Thr Lys Pro Phe Met Leu Pro Pro Val Ala Ala Ser Ser Leu
 180 185 190

Arg Asn Asp Ser Ser Ser Ser Asn Arg Lys Thr Asn Pro Ile Glu Asp
195 200 205

Ser Ser Ile Gln Trp Ala Val Met Ala Leu Pro Ala Cys Phe Ser Leu
210 215 220

Val Ile Gly Phe Ala Phe Gly Ala Phe Tyr Trp Lys Lys Lys Gln Pro
225 230 235 240

Asn Leu Thr Arg Thr Val Glu Asn Ile Gln Ile Asn Glu Glu Asp Asn
245 250 255

Glu Ile Ser Met Leu Gln Glu Lys Glu Arg Glu Phe Gln Glu Val
260 265 270

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 273 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

Met Lys Lys Thr Gln Thr Trp Ile Ile Thr Cys Ile Tyr Leu Gln Leu
1 5 10 15

Leu Leu Phe Asn Pro Leu Val His Thr Gln Gly Ile Cys Ser Asn Arg
20 25 30

Val Thr Asp Asp Val Lys Asp Val Thr Lys Leu Val Ala Asn Leu Pro
35 40 45

Lys Asp Tyr Met Ile Thr Leu Lys Tyr Val Pro Gly Met Asp Val Leu
50 55 60

Pro Ser His Cys Trp Ile Ser Glu Met Val Glu Gln Leu Ser Val Ser
65 70 75 80

Leu Thr Asp Leu Leu Asp Lys Phe Ser Asn Ile Ser Glu Gly Leu Ser
85 90 95

Asn Tyr Cys Ile Ile Asp Lys Leu Val Lys Ile Val Asp Asp Leu Val
100 105 110

Glu Cys Met Glu His Ser Ser Glu Asn Val Lys Lys Ser Ser Lys Ser
115 120 125

Pro Glu Pro Arg Gln Phe Thr Pro Glu Lys Phe Phe Gly Ile Phe Asn
130 135 140

Lys Ser Ile Asp Ala Phe Lys Asp Leu Glu Ile Val Ala Ser Lys Met
145 150 155 160

Ser Glu Cys Val Ile Ser Ser Thr Ser Ser Pro Glu Lys Asp Ser Arg
165 170 175

Val Ser Val Thr Lys Pro Phe Met Leu Pro Pro Val Ala Ala Ser Ser
180 185 190

Leu Arg Asn Asp Ser Ser Ser Ser Asn Arg Lys Ala Ser Asn Ser Ile
 195 200 205
 Glu Asp Ser Ser Leu Gln Trp Ala Ala Val Ala Leu Pro Ala Phe Phe
 210 215 220
 Ser Leu Val Ile Gly Phe Ala Phe Gly Ala Phe Tyr Trp Lys Lys Lys
 225 230 235 240
 Gln Pro Asn Leu Thr Arg Thr Val Glu Asn Arg Gln Ile Asn Glu Glu
 245 250 255
 Asp Asn Glu Ile Ser Met Leu Gln Glu Lys Glu Arg Glu Phe Gln Glu
 260 265 270
 Val

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 273 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

Met Lys Lys Thr Gln Thr Trp Ile Ile Thr Cys Ile Tyr Leu Gln Leu
 1 5 10 15
 Leu Leu Phe Asn Pro Leu Val Lys Thr Gln Glu Ile Cys Arg Asn Pro
 20 25 30
 Val Thr Asp Asn Val Lys Asp Ile Thr Lys Leu Val Ala Asn Leu Pro
 35 40 45
 Asn Asp Tyr Met Ile Thr Leu Asn Tyr Val Ala Gly Met Asp Val Leu
 50 55 60
 Pro Ser His Cys Trp Leu Arg Asp Met Val Thr His Leu Ser Val Ser
 65 70 75 80
 Leu Thr Thr Leu Leu Asp Lys Phe Ser Asn Ile Ser Glu Gly Leu Ser
 85 90 95
 Asn Tyr Ser Ile Ile Asp Lys Leu Gly Lys Ile Val Asp Asp Leu Val
 100 105 110
 Ala Cys Met Glu Glu Asn Ala Pro Leu Asn Val Lys Glu Ser Leu Lys
 115 120 125
 Lys Pro Glu Thr Arg Asn Phe Thr Pro Glu Glu Phe Phe Ser Ile Phe
 130 135 140
 Asn Arg Ser Ile Asp Ala Phe Lys Asp Phe Met Val Ala Ser Asp Thr
 145 150 155 160
 Ser Asp Cys Val Leu Ser Ser Thr Leu Gly Pro Glu Lys Asp Ser Arg
 165 170 175

Val Ser Val Thr Lys Pro Phe Met Leu Pro Pro Val Ala Ala Ser Ser
180 185 190

Leu Arg Asn Asp Ser Ser Ser Ser Asn Arg Lys Ala Ala Lys Ser Pro
195 200 205

Glu Asp Pro Gly Leu Gln Trp Thr Ala Met Ala Leu Pro Ala Leu Ile
210 215 220

Ser Leu Val Ile Gly Phe Ala Phe Gly Ala Leu Tyr Trp Lys Lys Lys
225 230 235 240

Gln Ser Ser Leu Thr Arg Ala Val Glu Asn Ile Gln Ile Asn Glu Glu
245 250 255

Asp Asn Glu Ile Ser Met Leu Gln Gln Lys Glu Arg Glu Phe Gln Glu
260 265 270

Val

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 273 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

Met Lys Lys Thr Gln Thr Trp Ile Ile Thr Cys Ile Tyr Leu Gln Leu
1 5 10 15

Leu Leu Phe Asn Pro Leu Val Lys Thr Lys Glu Ile Cys Gly Asn Pro
20 25 30

Val Thr Asp Asn Val Lys Asp Ile Thr Lys Leu Val Ala Asn Leu Pro
35 40 45

Asn Asp Tyr Met Ile Thr Leu Asn Tyr Val Ala Gly Met Asp Val Leu
50 55 60

Pro Ser His Cys Trp Leu Arg Asp Met Val Ile Gln Leu Ser Leu Ser
65 70 75 80

Leu Thr Thr Leu Leu Asp Lys Phe Ser Asn Ile Ser Glu Gly Leu Ser
85 90 95

Asn Tyr Ser Ile Ile Asp Lys Leu Gly Lys Ile Val Asp Asp Leu Val
100 105 110

Leu Cys Met Glu Glu Asn Ala Pro Lys Asn Ile Lys Glu Ser Pro Lys
115 120 125

Arg Pro Glu Thr Arg Ser Phe Thr Pro Glu Glu Phe Phe Ser Ile Phe
130 135 140

Asn Arg Ser Ile Asp Ala Phe Lys Asp Phe Met Val Ala Ser Asp Thr
145 150 155 160

Ser Asp Cys Val Leu Ser Ser Thr Leu Gly Pro Glu Lys Asp Ser Arg
 165 170 175
 Val Ser Val Thr Lys Pro Phe Met Leu Pro Pro Val Ala Ala Ser Ser
 180 185 190
 Leu Arg Asn Asp Ser Ser Ser Ser Asn Arg Lys Ala Ala Lys Ala Pro
 195 200 205
 Glu Asp Ser Gly Leu Gln Trp Thr Ala Met Ala Leu Pro Ala Leu Ile
 210 215 220
 Ser Leu Val Ile Gly Phe Ala Phe Gly Ala Leu Tyr Trp Lys Lys Lys
 225 230 235 240
 Gln Ser Ser Leu Thr Arg Ala Val Glu Asn Ile Gln Ile Asn Glu Glu
 245 250 255
 Asp Asn Glu Ile Ser Met Leu Gln Gln Lys Glu Arg Glu Phe Gln Glu
 260 265 270

Val

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 282 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

Thr Trp Ile Ile Thr Cys Phe Cys Leu Gln Leu Leu Leu Leu Asn Pro
 1 5 10 15
 Leu Val Lys Ala Gln Ser Ser Cys Gly Asn Pro Val Thr Asp Asp Val
 20 25 30
 Asn Asp Ile Ala Lys Leu Val Gly Asn Leu Pro Asn Asp Tyr Leu Ile
 35 40 45
 Thr Leu Lys Tyr Val Pro Lys Met Asp Ser Leu Pro Asn His Cys Trp
 50 55 60
 Leu His Leu Met Val Pro Glu Phe Ser Arg Ser Leu His Asn Leu Leu
 65 70 75 80
 Gln Lys Phe Ser Asp Ile Ser Asp Met Ser Asp Val Leu Ser Asn Tyr
 85 90 95
 Ser Ile Ile Asn Asn Leu Thr Arg Ile Ile Asn Asp Leu Met Ala Cys
 100 105 110
 Leu Ala Phe Asp Lys Asn Lys Asp Phe Ile Lys Glu Asn Gly Leu His
 115 120 125
 Tyr Glu Glu Asp Arg Phe Ile Pro Glu Asn Phe Phe Arg Leu Phe Asn
 130 135 140

Ser Thr Ile Glu Val Tyr Lys Glu Phe Ala Asp Ser Leu Asp Lys Asn
 145 150 155 160
 Asp Cys Ile Met Pro Ser Thr Val Glu Thr Pro Glu Asn Asp Ser Arg
 165 170 175
 Val Ala Val Thr Lys Thr Ile Ser Phe Pro Pro Val Ala Ala Ser Ser
 180 185 190
 Leu Arg Asn Asp Ser Ile Gly Ser Asn Thr Ser Ser Asn Ser Asn Lys
 195 200 205
 Glu Ala Leu Gly Phe Ile Ser Ser Ser Ser Leu Gln Gly Ile Ser Ile
 210 215 220
 Ala Leu Thr Ser Leu Leu Ser Leu Leu Ile Gly Phe Ile Leu Gly Ala
 225 230 235 240
 Ile Tyr Trp Lys Lys Thr His Pro Lys Ser Arg Pro Glu Ser Asn Glu
 245 250 255
 Thr Ile Gln Cys His Gly Cys Gln Glu Glu Asn Glu Ile Ser Met Leu
 260 265 270
 Gln Gln Lys Glu Lys Glu His Leu Gln Val
 275 280

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 266 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

Met Lys Lys Thr Gln Thr Trp Ile Ile Thr Cys Ile Tyr Leu Gln Leu
 1 5 10 15
 Leu Leu Phe Asn Pro Leu Val Lys Thr Gly Ile Cys Arg Asn Arg Val
 20 25 30
 Thr Asp Val Lys Asp Val Thr Lys Leu Val Ala Asn Leu Pro Lys Asp
 35 40 45
 Tyr Met Ile Thr Leu Lys Tyr Val Pro Gly Met Asp Val Leu Pro Ser
 50 55 60
 His Cys Trp Ile Ser Glu Met Val Glu Gln Leu Ser Val Ser Leu Thr
 65 70 75 80
 Asp Leu Leu Asp Lys Phe Ser Asn Ile Ser Glu Gly Leu Ser Asn Tyr
 85 90 95
 Ser Ile Ile Asp Lys Leu Val Lys Ile Val Asp Asp Leu Val Glu Cys
 100 105 110
 Glu Glu Asn Ser Ser Lys Asn Val Lys Lys Ser Lys Ser Pro Glu Pro
 115 120 125

Arg Leu Phe Thr Pro Glu Glu Phe Phe Arg Ile Phe Asn Arg Ser Ile
 130 135 140
 Asp Ala Phe Lys Asp Phe Met Val Ala Ser Lys Thr Ser Asp Cys Val
 145 150 155 160
 Val Ser Ser Thr Leu Ser Pro Glu Lys Asp Ser Arg Val Ser Val Thr
 165 170 175
 Lys Pro Phe Met Leu Pro Pro Val Ala Ala Ser Ser Leu Arg Asn Asp
 180 185 190
 Ser Ser Ser Ser Asn Arg Lys Ala Asn Glu Asp Ser Ser Leu Gln Trp
 195 200 205
 Ala Ala Met Ala Leu Pro Ala Leu Phe Ser Leu Val Ile Gly Phe Ala
 210 215 220
 Phe Gly Ala Leu Tyr Trp Lys Lys Lys Gln Pro Ser Leu Thr Arg Ala
 225 230 235 240
 Val Glu Asn Ile Gln Ile Asn Glu Glu Asp Asn Glu Ile Ser Met Leu
 245 250 255
 Gln Glu Lys Glu Arg Glu Phe Gln Glu Val
 260 265

(2) INFORMATION FOR SEQ ID NO:58:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 269 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

- (ix) FEATURE:
- (A) NAME/KEY: CDS
 - (B) LOCATION: join(1..210, 223..258)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

| | |
|---|-----|
| GAA TTC TTC CGT ATC TTC AAC CGT TCC ATC GAC GCT TTC AAA GAC TTC | 48 |
| Glu Phe Phe Arg Ile Phe Asn Arg Ser Ile Asp Ala Phe Lys Asp Phe | |
| 1 5 10 15 | |
| GTT GTT GCT TCC GAA ACC TCC GAC TGC GTT GTT TCC TCC ACC CTG TCT | 96 |
| Val Val Ala Ser Glu Thr Ser Asp Cys Val Val Ser Ser Thr Leu Ser | |
| 20 25 30 | |
| CCG GAA AAA GAC TCC CGT GTT TCG GTT ACC AAA CCG TTC ATG CTG CCG | 144 |
| Pro Glu Lys Asp Ser Arg Val Ser Val Thr Lys Pro Phe Met Leu Pro | |
| 35 40 45 | |
| CCG GTT GCT GCT TCC TCC CTG CGT AAC GAC TCC TCC TCC TCC AAC TCC | 192 |
| Pro Val Ala Ala Ser Ser Leu Arg Asn Asp Ser Ser Ser Ser Asn Ser | |
| 50 55 60 | |
| AAA TAC ATC TAC CTG ATC TAATAGGATC CG GTT ACC AAA CCG TTC ATG | 240 |
| Lys Tyr Ile Tyr Leu Ile Val Thr Lys Pro Phe Met | |
| 65 70 75 | |

CTG CCG CCG GTT GCT GCT TAATAGGATC C
 Leu Pro Pro Val Ala Ala
 80

269

(2) INFORMATION FOR SEQ ID NO:59:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 82 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

Glu Phe Phe Arg Ile Phe Asn Arg Ser Ile Asp Ala Phe Lys Asp Phe
 1 5 10 15
 Val Val Ala Ser Glu Thr Ser Asp Cys Val Val Ser Ser Thr Leu Ser
 20 25 30
 Pro Glu Lys Asp Ser Arg Val Ser Val Thr Lys Pro Phe Met Leu Pro
 35 40 45
 Pro Val Ala Ala Ser Ser Leu Arg Asn Asp Ser Ser Ser Ser Asn Ser
 50 55 60
 Lys Tyr Ile Tyr Leu Ile Val Thr Lys Pro Phe Met Leu Pro Pro Val
 65 70 75 80
 Ala Ala

(2) INFORMATION FOR SEQ ID NO:60:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1404 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

- (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 184..1002

- (ix) FEATURE:
 (A) NAME/KEY: mat_peptide
 (B) LOCATION: 259..1002

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

CCGCCTCGCG CCGAGACTAG AAGCGCTGCG GGAAGCAGGG ACAGTGGAGA GGGCGCTGCG 60
 CTCGGGCTAC CCAATGCGTG GACTATCTGC CGCCGCTGTT CGTGCAATAT GCTGGAGCTC 120
 CAGAACAGCT AAACGGAGTC GCCACACCAC TGTTTGTGCT GGATCGCAGC GCTGCCTTTC 180
 CTT ATG AAG AAG ACA CAA ACT TGG ATT CTC ACT TGC ATT TAT CTT CAG 228
 Met Lys Lys Thr Gln Thr Trp Ile Leu Thr Cys Ile Tyr Leu Gln
 -25 -20 -15

| | |
|---|-----|
| CTG CTC CTA TTT AAT CCT CTC GTC AAA ACT GAA GGG ATC TGC AGG AAT Leu Leu Leu Phe Asn Pro Leu Val Lys Thr Glu Gly Ile Cys Arg Asn -10 -5 1 5 | 276 |
| CGT GTG ACT AAT AAT GTA AAA GAC GTC ACT AAA TTG GTG GCA AAT CTT Arg Val Thr Asn Asn Val Lys Asp Val Thr Lys Leu Val Ala Asn Leu 10 15 20 | 324 |
| CCA AAA GAC TAC ATG ATA ACC CTC AAA TAT GTC CCC GGG ATG GAT GTT Pro Lys Asp Tyr Met Ile Thr Leu Lys Tyr Val Pro Gly Met Asp Val 25 30 35 | 372 |
| TTG CCA AGT CAT TGT TGG ATA AGC GAG ATG GTA GTA CAA TTG TCA GAC Leu Pro Ser His Cys Trp Ile Ser Glu Met Val Val Gln Leu Ser Asp 40 45 50 | 420 |
| AGC TTG ACT GAT CTT CTG GAC AAG TTT TCA AAT ATT TCT GAA GGC TTG Ser Leu Thr Asp Leu Leu Asp Lys Phe Ser Asn Ile Ser Glu Gly Leu 55 60 65 70 | 468 |
| AGT AAT TAT TCC ATC ATA GAC AAA CTT GTG AAT ATA GTG GAT GAC CTT Ser Asn Tyr Ser Ile Ile Asp Lys Leu Val Asn Ile Val Asp Asp Leu 75 80 85 | 516 |
| GTG GAG TGC GTG AAA GAA AAC TCA TCT AAG GAT CTA AAA AAA TCA TTC Val Glu Cys Val Lys Glu Asn Ser Ser Lys Asp Leu Lys Lys Ser Phe 90 95 100 | 564 |
| AAG AGC CCA GAA CCC AGG CTC TTT ACT CCT GAA GAA TTC TTT AGA ATT Lys Ser Pro Glu Pro Arg Leu Phe Thr Pro Glu Glu Phe Phe Arg Ile 105 110 115 | 612 |
| TTT AAT AGA TCC ATT GAT GCC TTC AAG GAC TTT GTA GTG GCA TCT GAA Phe Asn Arg Ser Ile Asp Ala Phe Lys Asp Phe Val Val Ala Ser Glu 120 125 130 | 660 |
| ACT AGT GAT TGT GTG GTT TCT TCA ACA TTA AGT CCT GAG AAA GAT TCC Thr Ser Asp Cys Val Val Ser Ser Thr Leu Ser Pro Glu Lys Asp Ser 135 140 145 150 | 708 |
| AGA GTC AGT GTC ACA AAA CCA TTT ATG TTA CCC CCT GTT GCA GCC AGC Arg Val Ser Val Thr Lys Pro Phe Met Leu Pro Pro Val Ala Ala Ser 155 160 165 | 756 |
| TCC CTT AGG AAT GAC AGC AGT AGC AGT AAT AGG AAG GCC AAA AAT CCC Ser Leu Arg Asn Asp Ser Ser Ser Ser Asn Arg Lys Ala Lys Asn Pro 170 175 180 | 804 |
| CCT GGA GAC TCC AGC CTA CAC TGG GCA GCC ATG GCA TTG CCA GCA TTG Pro Gly Asp Ser Ser Leu His Trp Ala Ala Met Ala Leu Pro Ala Leu 185 190 195 | 852 |
| TTT TCT CTT ATA ATT GGC TTT GCT TTT GGA GCC TTA TAC TGG AAG AAG Phe Ser Leu Ile Ile Gly Phe Ala Phe Gly Ala Leu Tyr Trp Lys Lys 200 205 210 | 900 |
| AGA CAG CCA AGT CTT ACA AGG GCA GTT GAA AAT ATA CAA ATT AAT GAA Arg Gln Pro Ser Leu Thr Arg Ala Val Glu Asn Ile Gln Ile Asn Glu 215 220 225 230 | 948 |
| GAG GAT AAT GAG ATA AGT ATG TTG CAA GAG AAA GAG AGA GAG TTT CAA Glu Asp Asn Glu Ile Ser Met Leu Gln Glu Lys Glu Arg Glu Phe Gln 235 240 245 | 996 |

GAA GTG TAATTGTGGC TTGTATCAAC ACTGTTACTT TCGTACATTG GCTGGTAACA 1052
Glu Val

GTTTCATGTTT GCTTCATAAA TGAAGCAGCT TTAACAAAT TCATATTCTG TCTGGAGTGA 1112
CAGACCACAT CTTTATCTGT TCTTGCTACC CATGACTTTA TATGGATGAT TCAGAAATTG 1172
GAACAGAATG TTTTACTGTG AACTGGCAC TGAATTAATC ATCTATAAAG AAGAACTTGC 1232
ATGGAGCAGG ACTCTATTTT AAGGACTGCG GGACTTGGGT CTCATTTAGA ACTTGCAGCT 1292
GATGTTGGAA GAGAAAGCAC GTGTCTCAGA CTGCATGTAC CATTTCATG GCTCCAGAAA 1352
TGTCTAAATG CTGAAAAAAC ACCTAGCTTT ATTCTTCAGA TACAACTGC AG 1404

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 273 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Lys | Lys | Thr | Gln | Thr | Trp | Ile | Leu | Thr | Cys | Ile | Tyr | Leu | Gln | Leu | -25 | -20 | -15 | -10 |
| Leu | Leu | Phe | Asn | Pro | Leu | Val | Lys | Thr | Glu | Gly | Ile | Cys | Arg | Asn | Arg | -5 | 1 | 5 | |
| Val | Thr | Asn | Asn | Val | Lys | Asp | Val | Thr | Lys | Leu | Val | Ala | Asn | Leu | Pro | 10 | 15 | 20 | |
| Lys | Asp | Tyr | Met | Ile | Thr | Leu | Lys | Tyr | Val | Pro | Gly | Met | Asp | Val | Leu | 25 | 30 | 35 | |
| Pro | Ser | His | Cys | Trp | Ile | Ser | Glu | Met | Val | Val | Gln | Leu | Ser | Asp | Ser | 40 | 45 | 50 | 55 |
| Leu | Thr | Asp | Leu | Leu | Asp | Lys | Phe | Ser | Asn | Ile | Ser | Glu | Gly | Leu | Ser | 60 | 65 | 70 | |
| Asn | Tyr | Ser | Ile | Ile | Asp | Lys | Leu | Val | Asn | Ile | Val | Asp | Asp | Leu | Val | 75 | 80 | 85 | |
| Glu | Cys | Val | Lys | Glu | Asn | Ser | Ser | Lys | Asp | Leu | Lys | Lys | Ser | Phe | Lys | 90 | 95 | 100 | |
| Ser | Pro | Glu | Pro | Arg | Leu | Phe | Thr | Pro | Glu | Glu | Phe | Phe | Arg | Ile | Phe | 105 | 110 | 115 | |
| Asn | Arg | Ser | Ile | Asp | Ala | Phe | Lys | Asp | Phe | Val | Val | Ala | Ser | Glu | Thr | 120 | 125 | 130 | 135 |
| Ser | Asp | Cys | Val | Val | Ser | Ser | Thr | Leu | Ser | Pro | Glu | Lys | Asp | Ser | Arg | 140 | 145 | 150 | |
| Val | Ser | Val | Thr | Lys | Pro | Phe | Met | Leu | Pro | Pro | Val | Ala | Ala | Ser | Ser | 155 | 160 | 165 | |

Leu Arg Asn Asp Ser Ser Ser Ser Asn Arg Lys Ala Lys Asn Pro Pro
 170 175 180
 Gly Asp Ser Ser Leu His Trp Ala Ala Met Ala Leu Pro Ala Leu Phe
 185 190 195
 Ser Leu Ile Ile Gly Phe Ala Phe Gly Ala Leu Tyr Trp Lys Lys Arg
 200 205 210 215
 Gln Pro Ser Leu Thr Arg Ala Val Glu Asn Ile Gln Ile Asn Glu Glu
 220 225 230
 Asp Asn Glu Ile Ser Met Leu Gln Glu Lys Glu Arg Glu Phe Gln Glu
 235 240 245

Val

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1088 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 151..885

(ix) FEATURE:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION: 226..885

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

AGCAGGGACA GTGGAGAGGG CGCTGCGCTC GGGCTACCCA ATGCGTGGAC TATCTGCCGC 60
 CGCTGTTCGT GCAATATGCT GGAGCTCCAG AACAGCTAAA CGGAGTCGCC ACACCACTGT 120
 TTGTGCTGGA TCGCAGCGCT GCCTTTCCTT ATG AAG AAG ACA CAA ACT TGG ATT 174
 Met Lys Lys Thr Gln Thr Trp Ile
 -25 -20
 CTC ACT TGC ATT TAT CTT CAG CTG CTC CTA TTT AAT CCT CTC GTC AAA 222
 Leu Thr Cys Ile Tyr Leu Gln Leu Leu Leu Phe Asn Pro Leu Val Lys
 -15 -10 -5
 ACT GAA GGG ATC TGC AGG AAT CGT GTG ACT AAT AAT GTA AAA GAC GTC 270
 Thr Glu Gly Ile Cys Arg Asn Arg Val Thr Asn Asn Val Lys Asp Val
 1 5 10 15
 ACT AAA TTG GTG GCA AAT CTT CCA AAA GAC TAC ATG ATA ACC CTC AAA 318
 Thr Lys Leu Val Ala Asn Leu Pro Lys Asp Tyr Met Ile Thr Leu Lys
 20 25 30
 TAT GTC CCC GGG ATG GAT GTT TTG CCA AGT CAT TGT TGG ATA AGC GAG 366
 Tyr Val Pro Gly Met Asp Val Leu Pro Ser His Cys Trp Ile Ser Glu
 35 40 45

| | |
|---|------|
| ATG GTA GTA CAA TTG TCA GAC AGC TTG ACT GAT CTT CTG GAC AAG TTT Met Val Val Gln Leu Ser Asp Ser Leu Thr Asp Leu Leu Asp Lys Phe 50 55 60 | 414 |
| TCA AAT ATT TCT GAA GGC TTG AGT AAT TAT TCC ATC ATA GAC AAA CTT Ser Asn Ile Ser Glu Gly Leu Ser Asn Tyr Ser Ile Ile Asp Lys Leu 65 70 75 | 462 |
| GTG AAT ATA GTG GAT GAC CTT GTG GAG TGC GTG AAA GAA AAC TCA TCT Val Asn Ile Val Asp Asp Leu Val Glu Cys Val Lys Glu Asn Ser Ser 80 85 90 95 | 510 |
| AAG GAT CTA AAA AAA TCA TTC AAG AGC CCA GAA CCC AGG CTC TTT ACT Lys Asp Leu Lys Lys Ser Phe Lys Ser Pro Glu Pro Arg Leu Phe Thr 100 105 110 | 558 |
| CCT GAA GAA TTC TTT AGA ATT TTT AAT AGA TCC ATT GAT GCC TTC AAG Pro Glu Glu Phe Phe Arg Ile Phe Asn Arg Ser Ile Asp Ala Phe Lys 115 120 125 | 606 |
| GAC TTT GTA GTG GCA TCT GAA ACT AGT GAT TGT GTG GTT TCT TCA ACA Asp Phe Val Val Ala Ser Glu Thr Ser Asp Cys Val Val Ser Ser Thr 130 135 140 | 654 |
| TTA AGT CCT GAG AAA GGG AAG GCC AAA AAT CCC CCT GGA GAC TCC AGC Leu Ser Pro Glu Lys Gly Lys Ala Lys Asn Pro Pro Gly Asp Ser Ser 145 150 155 | 702 |
| CTA CAC TGG GCA GCC ATG GCA TTG CCA GCA TTG TTT TCT CTT ATA ATT Leu His Trp Ala Ala Met Ala Leu Pro Ala Leu Phe Ser Leu Ile Ile 160 165 170 175 | 750 |
| GGC TTT GCT TTT GGA GCC TTA TAC TGG AAG AAG AGA CAG CCA AGT CTT Gly Phe Ala Phe Gly Ala Leu Tyr Trp Lys Lys Arg Gln Pro Ser Leu 180 185 190 | 798 |
| ACA AGG GCA GTT GAA AAT ATA CAA ATT AAT GAA GAG GAT AAT GAG ATA Thr Arg Ala Val Glu Asn Ile Gln Ile Asn Glu Glu Asp Asn Glu Ile 195 200 205 | 846 |
| AGT ATG TTG CAA GAG AAA GAG AGA GAG TTT CAA GAA GTG TAATTGTGGC Ser Met Leu Gln Glu Lys Glu Arg Glu Phe Gln Glu Val 210 215 220 | 895 |
| TTGTATCAAC ACTGTTACTT TCGTACATTG GCTGGTAACA GTTCATGTTT GCTTCATAAA | 955 |
| TGAAGCAGCT TTAACAAAT TCATATTCTG TCTGGAGTGA CAGACCACAT CTTTATCTGT | 1015 |
| TCTTGCTACC CATGACTTTA TATGGATGAT TCAGAAATTG GAACAGAATG TTTTACTGTG | 1075 |
| AAACTGGCAC TGA | 1088 |

(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 245 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

Met Lys Lys Thr Gln Thr Trp Ile Leu Thr Cys Ile Tyr Leu Gln Leu
 -25 -20 -15 -10
 Leu Leu Phe Asn Pro Leu Val Lys Thr Glu Gly Ile Cys Arg Asn Arg
 -5 1 5
 Val Thr Asn Asn Val Lys Asp Val Thr Lys Leu Val Ala Asn Leu Pro
 10 15 20
 Lys Asp Tyr Met Ile Thr Leu Lys Tyr Val Pro Gly Met Asp Val Leu
 25 30 35
 Pro Ser His Cys Trp Ile Ser Glu Met Val Val Gln Leu Ser Asp Ser
 40 45 50 55
 Leu Thr Asp Leu Leu Asp Lys Phe Ser Asn Ile Ser Glu Gly Leu Ser
 60 65 70
 Asn Tyr Ser Ile Ile Asp Lys Leu Val Asn Ile Val Asp Asp Leu Val
 75 80 85
 Glu Cys Val Lys Glu Asn Ser Ser Lys Asp Leu Lys Lys Ser Phe Lys
 90 95 100
 Ser Pro Glu Pro Arg Leu Phe Thr Pro Glu Glu Phe Phe Arg Ile Phe
 105 110 115
 Asn Arg Ser Ile Asp Ala Phe Lys Asp Phe Val Val Ala Ser Glu Thr
 120 125 130 135
 Ser Asp Cys Val Val Ser Ser Thr Leu Ser Pro Glu Lys Gly Lys Ala
 140 145 150
 Lys Asn Pro Pro Gly Asp Ser Ser Leu His Trp Ala Ala Met Ala Leu
 155 160 165
 Pro Ala Leu Phe Ser Leu Ile Ile Gly Phe Ala Phe Gly Ala Leu Tyr
 170 175 180
 Trp Lys Lys Arg Gln Pro Ser Leu Thr Arg Ala Val Glu Asn Ile Gln
 185 190 195
 Ile Asn Glu Glu Asp Asn Glu Ile Ser Met Leu Gln Glu Lys Glu Arg
 200 205 210 215
 Glu Phe Gln Glu Val
 220

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

Glu Glu Ile Cys Arg Asn Pro Val Thr Asp Asn Val Lys Asp Ile Thr
 1 5 10 15

Lys Leu Val Ala Asn Leu Pro Asn Asp Tyr Met Ile Thr Leu Asn Tyr
20 25 30
Val Ala Gly Met Asp Val Leu Pro Ser His Xaa Trp Leu Arg Asp
35 40 45

(2) INFORMATION FOR SEQ ID NO:65:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

Ile Thr Thr Leu Asn Tyr Val Ala Gly Met
1 5 10

(2) INFORMATION FOR SEQ ID NO:66:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

Val Ala Ser Asp Thr Ser Asp Cys Val Leu Ser Xaa Xaa Leu Gly Pro
1 5 10 15
Glu Lys Asp Ser Arg Val Ser Val Xaa Lys
20 25

(2) INFORMATION FOR SEQ ID NO:67:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 12 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

Asp Val Leu Pro Ser His Cys Trp Leu Arg Asp Met
1 5 10

(2) INFORMATION FOR SEQ ID NO:68:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 33 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

Glu Glu Asn Ala Pro Lys Asn Val Glu Ser Leu Lys Lys Pro Thr Arg
1 5 10 15

Asn Phe Thr Pro Glu Glu Phe Phe Ser Ile Phe Asp Arg Ser Ile Asp
20 25 30

Ala

(2) INFORMATION FOR SEQ ID NO:69:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

Glu Ser Leu Lys Lys Pro Glu Thr Arg
1 5

(2) INFORMATION FOR SEQ ID NO:70:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

Val Ser Val Xaa Lys
1 5

(2) INFORMATION FOR SEQ ID NO:71:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

Ile Val Asp Asp Leu Val Ala Ala Met Glu Glu Asn Ala Pro Lys
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:72:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

Asn Phe Thr Pro Glu Glu Phe Phe Ser Ile Phe Xaa Arg
1 5 10

(2) INFORMATION FOR SEQ ID NO:73:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

Leu Val Ala Asn Leu Pro Asn Asp Tyr Met Ile Thr Leu Asn Tyr Val
1 5 10 15

Ala Gly Asp Asp Val Leu Pro Ser His Cys Trp Leu Arg
20 25

(2) INFORMATION FOR SEQ ID NO:74:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

Ser Ile Asp Ala Phe Lys Asp Phe Met Val Ala Ser Asp Thr Ser Asp
1 5 10 15

Cys Val Leu Ser Xaa Xaa Leu Gly
20

(2) INFORMATION FOR SEQ ID NO:75:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

Glu Ser Leu Lys Lys Pro Glu Thr Arg Asn Phe Thr Pro Glu Glu Phe
1 5 10 15

Phe Ser Ile Phe Xaa Arg
20

(2) INFORMATION FOR SEQ ID NO:76:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 22 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

Glu Ser Leu Lys Lys Pro Glu Thr Arg Asn Phe Thr Pro Glu Glu Phe
1 5 10 15

Phe Ser Ile Phe Asp Arg
20

(2) INFORMATION FOR SEQ ID NO:77:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 8 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

Asn Ala Pro Lys Asn Val Lys Glu
1 5

(2) INFORMATION FOR SEQ ID NO:78:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 16 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

Ser Arg Val Ser Val Xaa Lys Pro Phe Met Leu Pro Pro Val Ala Ala
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:79:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 34 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

Ser Leu Lys Lys Pro Glu Thr Arg Asn Phe Thr Pro Glu Glu Phe Phe
1 5 10 15

Ser Ile Phe Asn Arg Ser Ile Asp Ala Phe Lys Asp Phe Met Val Ala
20 25 30

Ser Asp

(2) INFORMATION FOR SEQ ID NO:80:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

Ser Leu Lys Lys Pro Glu Thr Arg Asn Phe Thr Pro Glu Glu Phe Phe
1 5 10 15
Ser Ile Phe Asn Arg Ser Ile Asp Ala Phe Lys Asp Phe Met Val Ala
20 25 30
Ser Asp Thr Ser Asp
35

(2) INFORMATION FOR SEQ ID NO:81:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

Leu Arg Asp Met Val Thr His Leu Ser Val Ser Leu Thr Thr Leu Leu
1 5 10 15
Asp Lys Phe Ser Asn Ile Ser Glu Gly Leu Ser Asn Tyr Ser Ile Ile
20 25 30
Asp Lys Leu Gly Lys Ile Val Asp
35 40

(2) INFORMATION FOR SEQ ID NO:82:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

Ser Arg Val Ser Val Thr Lys Pro Phe Met Leu Pro Pro Val Ala Ala
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:83:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

Pro Val Ala Ala
1

(2) INFORMATION FOR SEQ ID NO:84:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

CCTGAGAAAG ATTCCAGAGTC

21

(2) INFORMATION FOR SEQ ID NO:85:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 19 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

CTGCAGTTTG TATCTGAAG

19

(2) INFORMATION FOR SEQ ID NO:86:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 19 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

CATATAAAGT CATGGGTAG

19

(2) INFORMATION FOR SEQ ID NO:87:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 27 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:
 ACTTGTGTCT TCTTCATAAG GAAAGCC 27

(2) INFORMATION FOR SEQ ID NO:88:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 21 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:
 TGTACGAAAG TAACAGTGTT G 21

(2) INFORMATION FOR SEQ ID NO:89:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 22 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:
 ACTGCTCCTA TTTAATCCTC TC 22

(2) INFORMATION FOR SEQ ID NO:90:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 23 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:
 CACTGACTCT GGAATCTTTC TCA 23

(2) INFORMATION FOR SEQ ID NO:91:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 15 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:
 TCGACCCGGA TCCCC 15

(2) INFORMATION FOR SEQ ID NO:92:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

TCGAGGGGAT CCGGG

15

(2) INFORMATION FOR SEQ ID NO:93:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 25 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:

TCTTCTTCAT GGC GCGGCA AGCTT

25

(2) INFORMATION FOR SEQ ID NO:94:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 18 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asp | Ser | Arg | Val | Ser | Val | Xaa | Lys | Pro | Phe | Phe | Met | Leu | Pro | Pro | Val |
| 1 | | | | 5 | | | | 10 | | | | | 15 | | |

Ala Ala

(2) INFORMATION FOR SEQ ID NO:95:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 18 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asp | Ser | Arg | Val | Ser | Val | Thr | Lys | Pro | Phe | Phe | Met | Leu | Pro | Pro | Val |
| 1 | | | | 5 | | | | 10 | | | | | 15 | | |

Ala Ala

(2) INFORMATION FOR SEQ ID NO:96:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 55 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:

CGATTTGATT CTAGAAGGAG GAATAACATA TGGTTAACGC GTTGAATTC GGTAC

55

(2) INFORMATION FOR SEQ ID NO:97:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 49 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

CGAATTCCAA CGCGTTAACC ATATGTTATT CCTCCTTCTA GAATCAAAT

49

(2) INFORMATION FOR SEQ ID NO:98:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 9 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

TATGCAGGA

9

(2) INFORMATION FOR SEQ ID NO:99:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 11 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

GATCTCCTGC A

11

(2) INFORMATION FOR SEQ ID NO:100:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 17 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:

TATGGAAGGT ATCTGCA

17

(2) INFORMATION FOR SEQ ID NO:101:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:

GATACCTTCC A

11

(2) INFORMATION FOR SEQ ID NO:102:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:

TTTCCTTATG

10

(2) INFORMATION FOR SEQ ID NO:103:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:

GCCGCCGCCA TG

12

(2) INFORMATION FOR SEQ ID NO:104:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:

TCTTCTTCAT GCGGCGGCA AGCTT

25